Where are the risks in high frequency trading?

Thierry FOUCAULT
Professor of Finance
HEC Paris

Progress in information and trading technologies have contributed to the development of high frequency traders (HFTs), that is, traders whose trading strategies rely on extremely fast reaction to market events. In this paper, the author describes HFTs' strategies and how they rely on speed. He then discusses how some of these strategies might create risks for financial markets. In particular, he emphasises the fact that extremely fast reaction to information can raise adverse selection costs and undermine incentives to produce information, reducing market participants' ability to share risks efficiently and asset price informativeness for resources allocation. The author also discusses recent extreme short-lived price dislocations in financial markets (e.g. the 2010 Flash crash) and argues that these events are more likely to be due to automation of trading and structural changes in market organisation rather than high frequency trading per se*. Throughout he argues that regulation of high frequency trading should target specific trading strategies rather than fast trading in general.*

A very important role of financial markets is to facilitate risk sharing among investors. To this end, the finance industry constantly innovates, by creating new financial instruments or new ways to trade (see Allen and Gale, 1994). Changes in trading technologies over the last thirty years offer a very good example. Trading has become increasingly automated, first in stock markets and more recently in derivatives, foreign exchange, and bond markets. Exchanges have replaced their trading floors by automated matching systems[1] and human traders (brokers or proprietary trading desks) are progressively replaced by machines and algorithms. This evolution also led to changes in the way information is disseminated to traders and gave rise to new forms of trading. In particular, automated trading allows extremely fast reaction to events and some trading firms' business models (so called high frequency traders) exploit this feature.

As other financial innovations, this evolution and high frequency trading raise many questions. For regulatory purposes, one must understand what are economic forces driving the growth of high frequency trading and their effects on the ability of financial markets to efficiently perform their functions (in particular risk sharing). In this paper, I discuss these points in light of recent academic findings regarding high frequency traders. My goal is not to provide an exhaustive survey of the quickly growing literature on this topic but rather to identify sources of risks associated with high frequency trading that deserve regulatory attention.[2]

1| Algorithmic and high frequency trading

Algorithmic trading encompasses a wide variety of strategies. For instance, brokers use algorithms to optimally slice and dice their orders over time and across different trading platforms (using so called smart routers) to reduce their price impacts and therefore execution costs for their clients. These strategies often call for frequent order submissions and cancellations, resulting in a sharp increase in the traffic on electronic trading platforms (see Chart 1).

Chart 1
The evolution of the quote-to-trade ratio
(nr. orders/nr. trades)

Source: Foucault, Kozhan, Tham (2015).
The figure shows the evolution of the quote-to-trade ratio around the introduction of Autoquote API on Reuters D-3000 (a trading platform in currency markets) in July 2003. Autoquote API allow computers to automatically enter orders on Reuters D-3000 without human intervention. Its introduction marks the beginning of algorithmic trading on Reuters D-3000.

Some proprietary trading firms' strategies rely on extremely fast reaction to market events, very broadly defined. For instance, a market event might be the arrival of news about a stock, a quote update for this stock, or a trade in assets with correlated payoffs (e.g. an option on the stock or a futures on a market index). To be fast, these firms invest in technologies that help them to minimise their trading "latencies", i.e. the time it takes for them to receive messages from data providers (e.g. trading platforms or data vendors such as Bloomberg, or Thomson-Reuters), process these messages, make a trading decision (e.g. the submission of a market order, a limit order, or the cancellation of orders previously submitted to the market), and finally implement this decision. For instance, they will invest in high speed connections to markets and data vendors, e.g. by buying the right to locate their servers in very close geographical proximity to trading platforms' own data servers (a practice known as "co-location") or by subscribing to direct data feeds to receive market data a split second before other market participants.[3]

These firms are usually called "high frequency traders" (HFTs). HFTs are one type of algorithmic traders because their strategies are computerised.

1 *For instance, the Paris Bourse switched to electronic trading in 1986 and the Chicago Mercantile Exchange closed its pit in July 2015.*
2 *See Biais and Foucault (2014) and SEC (2014) for more detailed surveys of the literature on high frequency trading.*
3 *For instance, in the United States, trading platforms must transmit their data to plan processors (the Consolidated Tape Association and Consolidated Quote Association), which consolidate the data and distribute them to the public. As this process takes a few milliseconds, market participants with direct access to the trading platforms' data feeds can obtain market data even faster than participants who obtain the data from plan sponsors (for a discussion, see SEC 2010, §IV.B.2). In Europe, there is yet no consolidated datafeed for stocks traded in multiple platforms.*

However, another defining characteristic of HFTs, not common to all algorithmic traders, is the very high speed at which they operate. For instance, using data on orders submitted by 15 HFTs on Nasdaq OMX Stockholm, Baron *et al.* (2015) find that the average minimum time elapsed between the submission of orders by the fastest traders in their sample is below one millisecond (of the order of one microsecond for the fastest trader). Well known independent trading firms with high frequency operations include KCG, Virtu, Flow traders, or Tradebot. Broker dealers and banks (such as Goldman Sachs, Morgan Stanley or Deutsche bank) or hedge funds (e.g. Citadel or Renaissance) also have high frequency trading desks.

There is so far no clear legal or regulatory definition of high frequency trading, which creates difficulties to analyse their effects on financial markets and regulation.[4] They are usually defined as being characterised (see SEC, 2010) by: (i) the placement of a large number of orders, (ii) the use of very high speed and algorithms to generate and execute their orders, (iii) the use of co-location services and individual data feeds provided by exchanges, (iv) the entry and exit of positions over very short time frames, (v) a high cancellation rate for their orders, and (vi) small end of the day positions.

Researchers rarely have access to datasets that "flag" orders with an identifier allowing them to distinguish orders placed by high frequency trading desks from orders placed by other market participants. Thus, they have often relied on indirect methods to identify those orders (see SEC, 2014 for a review of empirical studies on high frequency trading and the datasets used in these studies). Hence, one must be careful in interpreting existing empirical findings about high frequency trading. In particular, empirical regularities uncovered in these studies might in fact be due to strategies of participants that in fact are not HFTs.

Keeping this caveat in mind, estimates indicate that HFTs account for a significant share of the trading volume in electronic markets. For instance, for US equities markets, a report from the Tabb Group estimated that HFTs accounted for 51% of the number of shares traded in the United States. For twelve European trading platforms and 100 stocks, a study from the ESMA (2014) finds that pure HFT firms (i.e. excluding high frequency trading desks of investment banks) account for 24% of the value traded. HFTs are also present in foreign exchange markets, treasury markets, or commodities markets.

2| High frequency traders' trading strategies

The effects of HFTs on market quality are likely to depend on their trading strategies. Hence, before discussing these effects, it is useful to describe these strategies. They can be broadly classified in three categories: high frequency market making, high frequency arbitrage and high frequency directional trading.

2|1 High frequency market making

Market-makers post bid and ask quotes at which they stand ready to buy or sell shares of an asset. Thus, they are intermediaries between final sellers and buyers of an asset. For instance, a market maker might buy a stock from one investor at some point and then resell it after a while to another one. Alternatively, when the same asset is traded in multiple trading platforms (as in United States and European stock markets), a market maker can buy the asset on one platform (e.g. BATS in Europe) from one investor and resell it on another platform (e.g. Euronext).

Market makers are exposed to various risks (see Foucault, Pagano, and Röell, 2013): (i) the risk of fluctuations in the value of their positions ("inventory risk"), (ii) the risk of trading with better informed investors ("adverse selection risk") and (iii) the risk of trading at stale quotes when news arrives ("picking off risk"). Their bid-ask spread (the difference between the price at which they sell and the price at which they buy) is a compensation for these risks and therefore increases when they are higher. Bid-ask spreads are often used as measures of market illiquidity.

In principle, fast reaction to market events can alleviate some risks inherent to market making. First, by allowing market makers to turn around

4 *MIFID II defines high frequency trading as* "algorithmic trading that relies on computer program to determine the timing, prices or quantities of orders in fractions of a second."

their positions more quickly, speed can help them to reduce their inventory risk.[5] Second, speed allows market makers to update their quotes faster when news arrives, which reduces their exposure to the risk of trading at stale quotes.

Thus, speed can be a way for market makers to reduce their costs and thereby to post more competitive bid-ask spreads. In line with this idea, Brogaard *et al.* (2015) find that traders who subscribe to the fastest co-location service on Nasdaq OMX Stockholm have characteristics of market makers and that an upgrade in this service reduced their exposure to the risk of being picked off and their inventory costs.

2|2 High frequency arbitrage

Arbitrage opportunities between related assets are pervasive at the high frequency. For instance, consider an exchange traded fund (ETF) on a stock index. In theory, the price of the ETF must be equal to the value of the index (the value of the portfolio of constituent stocks of the index) at any point in time. If instead, the price of the ETF is above (below) the value of the index, an arbitrageur can immediately buy (sell) the portfolio of constituent stocks and sell (buy) the ETF, at a profit. In practice, such arbitrage opportunities appear frequently in ETF markets for two reasons. First, large buy or sell orders in the ETF (or constituent stocks) exert transient price pressures on the ETF, creating an arbitrage opportunity. Second, when information arrives, quotes for ETFs and constituent stocks are not updated at the same time (e.g. quotes in underlying stocks tend to be updated with a lag relative to quotes in the ETF market). This lack of perfect synchronisation in the adjustment of prices to information also give rise to arbitrage opportunities. The same type of opportunities arise more generally between derivatives assets (CDS, futures, options, etc.) and their underlyings, in currency markets (e.g. so called triangular arbitrage), between stocks traded on different platforms etc.[6]

These arbitrage opportunities are extremely short lived: they disappear as soon as market makers update their quotes or an arbitrageur exploit the opportunity. For instance, Budish *et al.* (2015) find that the median duration of arbitrage opportunities between the SPDR S&P 500 ETF (SPY) and the E-mini S&P 500 future from January 2005 to December 2011 varies from 250 milliseconds (in 2006) to about 10 milliseconds (in 2011). They also find on average 801 opportunities per day that deliver a potential profit of USD 98.01 per opportunity. Thus, a trader can exploit these small but frequent fleeting opportunities only if he is very fast. This has been another major impetus for the development of high frequency trading (see, for instance, Chaboud *et al.*, 2014 for a discussion in the context of currency markets).

2|3 High frequency directional trading

HFTs can also take position in anticipation of future price movements. This type of strategy is called "directional" because traders take a position in an asset in the direction of their expectation of a future price movement (e.g. they buy a stock if they expect its price to increase). Speed might be useful for this type of strategy because it enables traders to react faster to news.

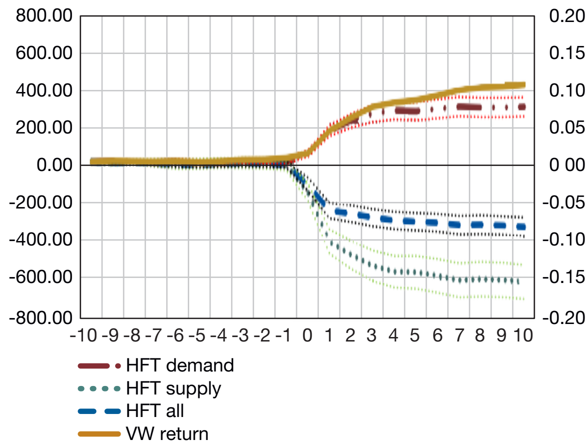
For instance, consider Chart 2 (taken from Brogaard *et al.*, 2014). The bold plain line shows the average price change over the ten seconds following the arrival (date 0) of "positive" (i.e. higher than expected) macroeconomic announcement for a sample of 120 Nasdaq stocks over the 2008-2009 period. On average, macroeconomic news in their sample moves prices by about 10 basis points in the 10 seconds following the news. Hence, a trader who reacts fast enough (say, in 10 milliseconds) to the release of positive macroeconomic news can expect to make a small profit on average by buying the stock. Brogaard *et al.* (2014) show that this is what some HFTs do. The dashed dotted line in Chart 2 shows the aggregate cumulative (over time) difference between

5 *For instance, consider a market maker with a long position in a French stock that trade on multiple markets (e.g. Euronext and BATS). If one particularly high bid price for the stock is posted on one market, a fast market maker can take advantage of this opportunity to unwind his position before other sellers take advantage of this bargain.*

6 *See, for instance, Ben-David* et al. *(2015) for evidence on arbitrage opportunities in ETF markets, or Chaboud* et al. *(2014) and Foucault, Kozhan and Tham (2015) for evidence of arbitrage opportunities in currency markets.*

Chart 2
Positive macro news

(x axis: seconds around macro news announcements)
HFT order flow USD 10,000 *Returns in %*



Source: Brogaard, Hendershott and Ryordan (2014).

"aggressive" purchases and sales (i.e. executed with marketable orders) by HFTs in Brogaard *et al.* (2014)'s sample. Similarly, the dotted line shows the aggregate cumulative difference between "passive" purchases and sales (i.e. executed with limit orders) by high frequency traders in this sample.[7]

Clearly, some high frequency traders quickly accumulate a significant long position (by placing buy market orders) just after the arrival of the positive macro news, presumably in anticipation of the price rise that materialises in the subsequent 10 seconds. Interestingly, other HFTs (those who trade with limit orders) sell shares just after the announcement, as the dotted line shows. The symmetry between the dashed dotted line and the dotted line suggests that in fact the latter do not have time to revise their quotes after the news arrival and get adversely executed ("picked off") by HFTs who are fast enough to take advantage of the macro announcement before prices adjust (see Dugast, 2015 for a model of price adjustments after news that formalises this scenario).

Macroeconomic announcements only constitute a small fraction of all "news" in a given day. In fact, progress in information technologies enable traders to react to a myriad of signals that in principle could move market prices. Accordingly, some information providers (such as Thomson-Reuters, Bloomberg, or Dataminr) now provide buy and sell signals extracted from raw information available in social medias such as tweeter.[8] Moreover, market data (quotes, trades, order submission etc.) themselves constitute a piece of information about future price movements.[9] Having access to these data faster is another way to anticipate price movements in the short run. Consistent with this idea, Brogaard *et al.* (2015) find that lagged one second returns of futures contracts on the OMXS30 index forecast the direction of market orders submitted by fast traders in their sample (30 Swedish stocks constituents of the OMXS30 index), suggesting that these traders use information in future returns to forecast impending price changes in constituent stocks (see also Zhang, 2012 for similar findings).

Another source of advantage for high frequency traders might stem from their ability to process vast amount of data quickly due to their massive investments in computers and efficient algorithms. This capacity might be useful in particular to better filter out noise from market data and thereby obtain more accurate signal to forecast future price movements. Foucault, Hombert, and Rosu (2016) derive the optimal trading strategy of an investor who reacts faster to news and filters out noise from news more efficiently than other market participants. They show that speed matters: the equilibrium trading strategy of the investor is significantly different from that of an investor who is just skilled at processing information (as in traditional models of trading with asymmetric information).

In addition, HFTs' computing power (and sophisticated data analysis techniques) might help

7 *A limit order is an order to buy or sell a given number of shares at a given price. In general, this price is such that the order cannot be filled immediately. In this case, a limit order is stored in a limit order book until another investor accepts to trade at its price or until the limit order submitter cancels his order. "Marketable orders" are orders to buy or sell a given number of shares at a price such that they can be filled upon submission against other limit orders standing in the limit order book. Trading platforms often refer to marketable orders as "aggressive" orders in the sense that their submission triggers a trade. Limit orders are "passive" in the sense that their execution can only be triggered by the arrival of a marketable order.*

8 *See "Mining for tweets of gold",* The Economist, *June 7, 2014, or "How investors are using social medias to make money",* Fortune, *December 7, 2015.*

9 *This is in fact consistent with theories such as Grossman and Stiglitz (1980) or Blume, Easley, and O'Hara (1994) that predict that market data such as stock prices or trading volume contain information that can be used to forecast future returns.*

them to detect footprints left by other traders when the latter execute large orders (see Hirshey, 2013 and van Kervel and Menkveld, 2015 for evidence). Indeed, such large orders are often split in a chain of smaller orders (called "child orders") to reduce their impact on prices. The detection of early child orders in this chain might then be useful to forecast the arrival of later child orders.

This "order anticipation" strategy might be profitable for at least two reasons. First, traders placing large orders might themselves be informed. Thus, their buys (sales) forecast a price increase in the future (decrease). In this case, it is optimal for order anticipators to mimick these trades (see Yang and Zhu, 2015 for a theoretical analysis). In this scenario, order anticipators compete away informed traders' profits. Second, traders placing large orders might be forced to liquidate a large position because of funding needs (e.g. an hedge fund might liquidate a large position to meet margin calls). Such forced liquidation by distressed traders generally occur at discounted prices. As shown by Brunnemeier and Pedersen (2005), traders who correctly infer the presence of such a distressed trader have an incentive to (i) initially trade in the same direction as the distressed trader to amplify the downward price pressure due to the distressed trader's orders and (ii) eventually buy the asset at a deep discounted price (Brunnermeier and Pedersen, 2005 refer to this strategy as "predatory trading").

Using data on HFTs' orders and institutional investors on Nasdaq OMX, Menkveld and van Kervel (2015) do not find evidence that HFTs in their sample engage in predatory trading around large institutional trades. Indeed, they appear to trade against early child orders from institutional investors, thereby dampening their impact on prices. However, HFTs eventually turn around their position and start trading in the same direction as early child orders. This behaviour might be consistent with either optimal risk management by HFTs or the "order anticipation hypothesis" according to which HFTs "mimick" institutional investors' informed trades, once they have inferred the presence of such investors from past trades.

High frequency traders have also been accused to engage in price manipulation. In particular, two strategies ("momentum ignition" and "spoofing") have attracted attention. "Momentum ignition," consists in placing buy market (or sell market orders) in the expectation that this behaviour will induce other traders to do the same. The flurry of buy orders that ensue might then push prices up, allowing the "momentum ignitor" to liquidate his position at a profit. "Spoofing" consists in entering, say, buy limit orders and cancelling them quickly in the hope that this will induce other traders to buy the asset and allow the manipulator to gain from execution of sell limit orders at inflated price. This practice is banned by the Dodd-Frank act and, recently, several traders have been charged of "spoofing" in the United States.[10]

Defining price or market manipulation is difficult, both in legal and economic terms (see Fishel and Ross, 1991 and Kyle and Viswnanathan, 2008). In particular, as Kyle and Viswanathan (2008) point out, it is difficult to distinguish between trading strategies that undermine both price informativeness and liquidity (which Kyle and Viwanathan, 2008 view as manipulative) from trading strategies that might look manipulative, but which just consist in rational exploitation of market power and private information. Several models show that the optimal behaviour of informed investors can be complex and counter-intuitive and yet their trades make prices more informative and do not intend to be manipulative.[11] The same problem arises for interpreting the intent of HFTs' order submission patterns.

Market making, arbitrage, directional trading, order anticipation, and manipulative strategies have a long history in financial markets and they have all been extensively analysed by economists. What is novel is the intensive use of information technologies to implement these strategies and the way they are implemented. On this, very little information is available because high frequency trading desks see this implementation as the source of their competitive advantage and naturally make all efforts to protect their "secret sauce."[12]

10 *See "Flash crash: trading terms and manipulation techniques explained",* Financial Times*, April 22, 2015 and "Regulators step up efforts to stop spoofing",* Financial Times*, November 5, 2015.*

11 *For instance, Back and Baruch (2004) show that randomising between buy and sell orders can be a way to minimise trading costs for an informed investor.*

12 *A good example is Alexei Aleynikov's case. He was charged of stealing high-frequency trading code from Goldman Sachs (see "Ex-Goldman programmer guilty of stealing code",* New York Times*, May 2015).*

As different strategies leave different footprints in the data, it is possible to infer – to a limited extent – HFTs' strategies. For instance, market makers tend to mainly use limit orders (so called passive orders) while directional traders who seek to profit from very short term price changes should predominantly use marketable orders (as limit orders take time to execute and might not execute at all). In a recent survey of the empirical literature, the SEC notes that: "*Perhaps the most noteworthy finding of the HFT Dataset papers is that HFT is not a monolithic phenomenon but rather encompasses a diverse range of strategies. In particular, HFT is not solely, or even primarily, characterized by passive market making strategies that employ liquidity providing orders […] Moreover, the level and nature of HFT activity can vary greatly across different types of stocks*." (SEC, 2014).

This observation is important because the effects of high frequency trading on market quality are more likely to depend on the type of strategies that fast traders use (see below) than speed *per se*. In practice large high frequency trading firms are likely to be opportunistic and use a strategy as long as it is profitable (and, hopefully, legal) and exit it when it becomes unprofitable. The data however suggests that there is some degree of specialisation among high frequency trading firms, some appearing to be more specialised in market making while others more specialised in arbitrage or directional trading (see Hagströmer and Norden, 2013).

3| High frequency trading: risks and benefits

High frequency trading has attracted a lot of media and regulatory attention, with claims from popular writers that high frequency trading could harm other market participants and threaten the integrity of financial markets (see, in particular, Michael Lewis, 2014). A key issue is whether high frequency trading enhances or undermines (i) market liquidity for risk sharing and (ii) pricing "accuracy", i.e. the informativeness of asset prices for resources allocation since risk sharing and information production are two important functions of financial markets. Another concern is that high frequency trading may jeopardise market stability. I discuss these points below.

3|1 Private vs. social benefits

As explained previously, HFTs massively invest in trading technologies and information. This suggests that they individually benefit from these investments. The social benefit of these investments is less clear, however. Indeed, they give an edge to HFTs in accessing to information and reacting fast to it, which is a source of adverse selection for other participants.

For instance, consider Chart 2 again. It shows that some HFTs place buy market orders just after the arrival of positive macroeconomic news and slightly in advance of a price increase due to these news. The gain made by these traders on their buys is a loss for their counterparties, who are "adversely selected" by better informed parties. More generally, empirical evidence on high frequency trading (see, for instance, Baron, Brogaard, and Kirilenko, 2014 or Brogaard, Hendershott, and Riordan, 2014) suggests that market orders from HFTs are informed (anticipate on future price movements) and thereby generate adverse selection costs for their counterparties (including HFTs specialising in liquidity supply). For example, Brogaard, Hendershott, and Riordan (2014) write (on page 2268): "*We show that HFTs impose adverse selection costs on other investors*."

One might argue that this is not a problem because fast trading on information is just a monetary transfer from fast to slow traders, i.e. a zero sum game. This redistribution can be perceived as unfair (for slow traders) but there is no welfare loss in aggregate. Biais, Foucault, and Moinas (2015) show that this argument is incomplete for two reasons. First, adverse selection implies that all traders bear larger impact costs when they trade. As the cost of trading gets larger, investors with small gains from trade (relative to the cost of trading) trade less (e.g. hedge risks less efficiently) or stop trading. Second, investment by HFTs must be accounted for in computing the social gains and benefits of this activity. If this activity just serves to play a zero sum game then its social cost is equal to the resources allocated to it. These resources are significant. For instance, the Project Express by Hibernia Atlantic drew a new fiber optic cable across the Atlantic, to increase by 5 millisecond the time to connect Wall Street to the City at a cost USD 300 million. Eventually, the cost of this project has to be covered by fees charged to firms using this cable to get fast access to information. For 2013 alone, the Tabb Group estimates the investment in fast trading technologies at USD 1.5 billion, twice the amount invested in 2012 (see *The Wall Street Journal*, 2014).

Not all HFTs are directional traders, however. As explained previously, some use their fast access to markets and information for market-making. If, in this case, trading speed reduces the cost of market making (inventory costs and the cost of being picked off by faster traders) or if it intensifies competition among market makers then HFTs should reduce transaction costs for investors. Moreover, if high frequency market-making reduces the cost of intermediation then this reduction is a social benefit.

It has been difficult so far to separately measure the effects of various type of trading strategies used by HFTs. Indeed, empiricists often observe HFTs' actions (order submissions, trades etc.), directly or indirectly, but not the strategies that command these actions. Yet these are the strategies that matter for observed effects. For instance, Brogaard *et al.* (2016) analyse the evolution of liquidity measures (e.g. the effective bid-ask spread) for 30 stocks on Nasdaq OMX around an up-grade in colocation services provided by Nasdaq OMX. They find that this up-grade results in an improvement in liquidity (smaller bid-ask spreads). Decomposition of this effect shows that the upgrade has two effects: (i) it results in smaller "realised bid-ask spreads" (a measure of per trade profit net of adverse selection costs for market makers) and (ii) larger price impacts (a measure of adverse selection costs borne by market makers). The decrease in realised spreads is consistent with investment in speed intensifying competition among market makers while the increase in price impacts is consistent with investment in speed increasing exposure to adverse selection for market makers.

The decrease in realised spreads more than offsets the increase in price impacts in Brogaard *et al.* (2016)'s sample so that, in net, investment in speed (co-location up-grade) appears beneficial. This analysis however shows the importance of measuring separately the effects of various strategies (maybe the benefit of co-location up-grades could have been even stronger if co-location could not be used to implement directional strategies).

Regulatory interventions about high frequency trading should therefore target specific trading strategies rather than high frequency trading in general. Indeed, regulation should not discourage high frequency trading when it serves to decrease trading costs for investors. In contrast, it should discourage strategies that exploit small differences in the speed of access to information about future price movements.

Consider for instance the proposal to tax traders when their order-to-trade ratio exceeds a certain threshold (as planned by MiFID II). Such a tax makes cancellations of their orders more costly for traders who mainly use limit orders, i.e. those who most likely are market makers. These traders need to frequently cancel their limit orders to (i) optimally control their inventory risk (cancellations can be part of an optimal inventory management strategy), (ii) reduce their exposure to the risk of trading at stale quotes, (iii) account for information contained in trades in other trading platforms (see van Kervel, 2015). If canceling orders become more costly, high frequency market makers must therefore raise their bid-ask spreads because their costs of market making increase. In contrast, directional traders (those trading on information) mainly submit market orders. Thus they naturally have smaller order-to-trade ratios. Hence, a tax on order-to-trade ratios is rather counterproductive: it discourages high frequency market making (which is likely to improve liquidity) while having no effect on directional HFTs (who harm liquidity). A more effective tool would be to add a very small random delay to the execution time for market orders. Indeed, this delay should reduce HFTs' ability to pick off stale quotes while leaving the possibility to traders submitting limit orders to revise their quotes.

3|2 Foreknowledge vs. discovery

In addition to facilitate risk sharing and the realisation of gains from trade, another important function of financial markets is to produce information. That is, asset prices aggregate informed investors' signals and thereby convey information for real decisions, e.g. investment (see Bond, Edmans, and Goldstein, 2012).[13] Thus, one potential benefit

13 *For instance, Fama and Miller (1972, p.335) write:* "An efficient market has a very desirable feature. In particular, at any point in time market prices of securities provide accurate signals for resource allocation; that is firms can make production-investment decisions."

of informed trading is that by making prices more informative, it also makes real decisions more efficient.

A natural conjecture is that by trading faster on information, high frequency directional traders accelerate the speed at which prices reflect information. Brogaard *et al.* (2014) find supportive evidence. They show that HFTs trade in the opposite direction to transitory pricing errors and in the same direction as permanent changes in asset value. In other words, HFTs' make prices closer to a random walk, as should be the case in an informationally efficient market. There is also evidence that the growth of algorithmic trading is associated with more short-lived arbitrage opportunities (see Budish *et al.,* 2015 for evidence about cross market arbitrage in ETFs and Chaboud *et al.,* 2014 for triangular arbitrage).

It does not follow however that HFTs make prices more *informative* as signals for resources allocation.[14] Hirshleifer (1971) distinguishes "foreknowledge," i.e. information about a state that, in due time, will be known to all, from "discovery," i.e. the production of information that would not be known without active human intervention. This distinction is very relevant to think about the effect of high frequency trading on price informativeness. Indeed, consider again the case of positive macroeconomic announcements in Chart 1. By observing macroeconomic news slightly faster than other traders, some HFTs obtain foreknowledge of an information that will be evident to all in a few seconds. By trading fast on this information, they accelerate the speed at which prices reflect this information, making markets more informationally efficient. However, they do not "discover" information: the macroeconomic announcement takes place whether or not some traders observe it very fast. In contrast, an analyst that combines various data about a firm to assess its value produces information that would not be available otherwise. By trading on this information, the analyst incorporates in prices information that would not otherwise be available.

There is no evidence so far that HFTs contribute to incorporate in asset prices information that otherwise would not be available. In fact, to the extent that HFTs reduces profits from discovering information, they might make asset prices less informative. For instance, Dugast and Foucault (2016) consider a model in which investors can choose to trade on raw information (e.g. buy or sell signals based on the content of tweets; see *Fortune*, December 7, 2015.) or on processed information (information processing is a form of discovery). Processed information is more accurate but processing information takes time. Thus, trading on raw information is a form of fast trading. They show that a reduction in the cost of access to raw information can discourage the production of processed information and thereby make prices less informative in the long run. Interestingly, Weller (2016) finds empirically a negative relationship between measures of algorithmic trading in US stocks and the extent to which prices contain information upon forthcoming earnings announcements. This suggests that the incentive to produce information about firms' fundamentals is smaller in stocks with a high level of algorithmic activity.

3|3 Fragility vs. reliability

In recent years, US financial markets have experienced several severe but short price disruptions. Three events are particularly noteworthy.

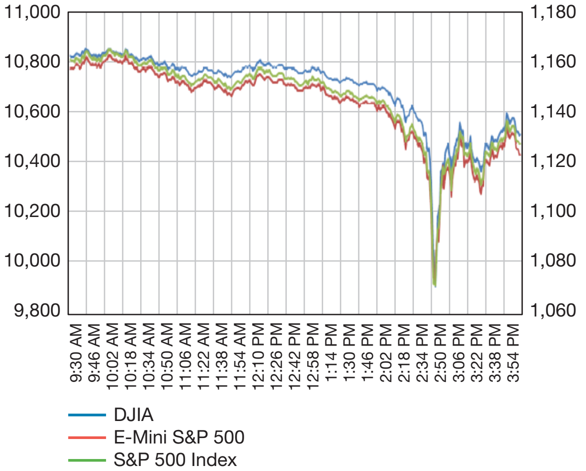
- **The flash crash of May 6, 2010 (see Chart 3)**. Over an interval of 30 minutes starting at 2:32 p.m. (Eastern Time), US stock indices, exchange traded funds, and index futures experienced very large upward and downward price movements. In particular, the Dow Jones lost about 6% in about five minutes, with stocks trading at very distorted prices (e.g. Accenture at one penny per share or Apple at USD 100,000 per share). The crash affected stock markets, exchange traded funds and futures markets. By 3:00 p.m., prices had reverted to levels close to their pre-crash level.

- **The Treasury flash crash of October 15, 2014 (see Chart 4)**. Between 9:33 and 9:40 a.m., the yield on the US 10-year Treasury bond fell by

14 *The reason is that the notion of informational efficiency is defined with respect to the "stock" of available information. It says nothing on amount of available information itself. If there is no information available, asset prices will be completely uninformative, even if they are informationally efficient.*

Chart 3
End-of-minute transaction prices of the Dow Jones Industrial Average (DJIA), S&P 500 Index, and the June 2010 E-Mini S&P 500 futures

(contract on May 6, 2010 between 8:30 and 15:15 CT)



Source: CFTC/SEC Staff Report, "Preliminary Findings Regarding the Market Events of May 6, 2010", 2010.
Note: This figure presents end-of-minute transaction prices of the Dow Jones Industrial Average (DJIA), S&P 500 Index, and the June 2010 E-Mini S&P 500 futures contract on May 6, 2010 between 8:30 and 15:15 CT.

Chart 4
10-year Treasury yield in cash market on October 15

Source: US Department of the Treasury, Board of Governors of the Federal Reserve System, Federal Reserve Bank of New York, SEC and CFTC, Joint Staff Report (2015), Chart 2.1.

16 basis points. The entire Treasury bonds and futures yield curve was affected as were, to a lesser extent, interest swaps and equity markets. By 10:00 a.m., yields were back to their level before the outset of the crash.

- **The exchange traded fund (ETFs) flash crash of August 24, 2015**. At the opening of the market, at 9:30 a.m. on August 24, 2015, the price of several exchange traded funds in the United States declined significantly relative to the indices that they track. For instance, the SPDR S&P 500 ETF (SPY) opened for regular trading hours at a discount of 5.2% relative to its previous day closing. This discount deepened further by 7.8% by 9:35. The SPY price then quickly reverted above the opening price. The drop in price for the SPY relative to the previous day price was one of the largest in the last decade (see SEC, 2015). The 50 largest exchange traded products (about 40% of all these products) experienced a decline in prices by more than 10% (SEC, 2015). Moreover, from 9:30 a.m. to 9:45 a.m., a large number of large capitalisation stocks on the NYSE and Nasdaq experienced drop in prices larger than 10%.

All these events share some common features. First, the extreme price movements observed during these events are accompanied by a sharp decline in liquidity of the affected markets (see Joint Staff Report, 2015 for the Treasury flash crash and CFTC/SEC Staff Report, 2010 for the 2010 Flash crash). Thus, these crashes are both price and liquidity crashes. Second, they happened without apparent changes in fundamentals. In fact, in each case, the quick price reversal that follows the initial high drop or spike in prices suggests that these price movements are not due to a change in fundamentals. Third, multiple assets are affected, maybe due to spillovers effects between asset classes linked by no-arbitrage relationships.

In addition, market participants claim that "mini flash crashes" (sudden drop or spike in prices followed by a price reversal in few seconds in one asset) happen routinely in today's markets.[15] These flash crashes are labeled "mini" because they do not simultaneously

15 *For instance, according to an article from the* Huffington Post*: "[...]* mini-flash crashes happen all of the time now. Just Monday, shares of Google collapsed briefly in a barely noticed flash crash of one of the country's biggest and most important companies." *(See* Huffington Post*, 2004). Similarly, Nanex (a financial data provider) reports more than 18,000 mini flash-crashes from 2006 to 2010 in US equity markets, that is, about USD 195 per month (Nanex defines a flash-crash as an up or down price movement greater than 0.8% in less than 1.5 second). See* http://www.nanex.net/FlashCrashEquities/FlashCrashAnalysis_Equities.html

affect a wide number of assets, in contrast to the three crashes discussed above. Yet, these events are potentially problematic because they might potentially be a catalyst for wider market disruptions.

In the aftermath of the 2010 Flash crash, several commentators suggested that high frequency trading might have played a role in the crash and that high speed trading was making markets more fragile. For instance, in September 2010, speaking before the Security Traders Association, Mary Schapiro, then Chairman of the Securities and Exchange Commission (SEC), said: "*Given their volume and access, high frequency trading firms have a tremendous capacity to affect the stability and integrity of the equity markets. Currently, however, [they]…. are subject to very little in the way of obligations either to protect that stability…. in tough times, or to refrain from exacerbating price volatility…. An out-of-control algorithm…. can also cause severe trading disruptions that harm market stability and shake investor confidence*."

Yet, it is far from clear that HFTs played a role in flash crashes and detailed reports by public agencies on these events do not show or even suggest that HFTs were directly responsible for these events (see, for instance, SEC, 2015, Joint Staff Report, 2015 or CFTC/SEC Staff Report, 2010). In fact, the cause(s) of flash crashes (or mini flash crashes) and the mechanisms for propagation of shocks across asset classes are still far from being well understood. It is likely that these events are due to a combination of factors and that automation, rather than speed of trading alone, might have played a role.[16]

In fact, it is useful to recognise that the automation of trading creates new operational risks.[17] Consider first the growth of algorithmic trading. Several recent events show that errors in the design of these algorithms can be a source of serious failures in financial markets. For instance, the IPO of BATS Global market failed on March 23, 2012 because the algorithm used to run the electronic auction for this IPO went wrong.[18] Another example is the loss by Knight Capital (a major broker-dealer in the United States until 2012) of 440 million in 30 minutes on August 1, 2012 due to an ill-designed algorithm, leading to the acquisition of Knight by one of its competitor (GETCO).

Trading platforms' matching systems and data feed used to disseminate information on market conditions can also experience technical glitches. For instance, on July 9, 2015, a computer glitch (due to the rollout of a new software) led to a trading halt of two hours on the NYSE. Automation might also leverage people ability to engage in market manipulation or deliberate attempt to affect the integrity of financial markets (for terrorism purpose for instance). For example, in April 2015, an independent London-based trader (Mr. Navinder Singh Sarao) was charged by US authorities of directly causing the Flash crash using a "spoofing" strategy (see Section 2).[19] Another example is "quote stuffing", a strategy that consists in deliberately increasing the number of messages (e.g. limit order submission followed by cancellations) sent to a trading platform to slow down other traders (e.g. the speed at which they receive information from exchanges' datafeed).[20]

The race for quick access to trading platforms or market data can also be a source of operational risk when it leads to traders to bypass safety checks. For instance, for risk management purpose, trading firms usually embed in their algorithms automated checks of their orders, to guarantee that they do not expose the firm to too large risks. Yet, these checks take time and therefore conflict with economic incentives to be fast for HFTs. Similarly, brokers usually impose automated risks limits on their clients. However, high frequency trading firms often bypass these automated limits by requiring so called direct market access (DMA) to again increase the speed at which they can submit orders to markets.

Given the automation and increasing complexity of financial markets, it should not be surprising that technological mishaps sometimes happen as it is

16 *Other possible factors might be market fragmentation (see Madhavan, 2012) or changes in the nature of liquidity provision, in particular the disappearance of designated market makers.*
17 *The Basel Committee on Banking Supervision (2001) defines "operational risk" as:* "The risk of loss resulting from inadequate or failed internal processes, people, and systems or from external event."
18 *Ironically, BATS operates one of the major electronic trading platform in the United States and in Europe.*
19 *The case is not settled. It is unclear whether Mr Sarao's orders were manipulative and unclear whether, if so, they could alone have triggered the Flash crash (see* Financial Times*, 2015).*
20 *See Ye, Yao, and Gai (2013) for evidence on quote stuffing.*

the case in other industries (e.g. transportation). Kumiega, Sterijevski, and van Vliet (2016) argue that one should therefore use notions from industrial engineering such as "reliability", which is an estimate of the probability of a mishap event, to evaluate the performance of financial markets and their fragility. They claim that from this perspective financial markets are on par with other industries. For instance, Gao and Mizrach (2015) have measured the frequency of large but transient price movements during the day (which they call "breakdowns" or "breakups") in US stocks.[21] They find that the frequency of such events has decreased since 2000 to less than 1% per stock-day in recent years, in contrast to the perception that they are more frequent. More empirical work is needed to assess the reliability of current market structures.

Several recent regulatory initiatives aim at reducing operational risks due to algorithmic trading (including high frequency trading) and automation. For instance, according to the SEC's Regulation Systems Compliance and Integrity rule (SEC, 2013), exchanges and traders using computerised trading systems must "*establish written policies and procedures reasonably designed to ensure that their systems have levels of capacity, integrity, resiliency, availability, and security adequate to maintain their operational capability and promote the maintenance of fair and orderly markets, and that they operate in the manner intended*" (p. 3). The rule would also require exchanges and traders to regularly test their systems and have disaster plans in place. MiFID II in Europe has qualitatively similar provisions for the algorithms used by HFTs. For instance, according to MiFID II, HFTs will have to develop effective systems and risk controls and to report their algorithmic strategies to regulators.

As mentioned previously, the May 2010 Flash crash or the Treasury flash crash were accompanied by a sudden evaporation of liquidity. This evaporation suggests that one source of fragility stems from a change in the nature of liquidity provision in electronic markets. In particular, one concern is that high frequency market makers do not have the ability to provide significant liquidity in times of market stress or might even withdraw from the market in such times. Again, academic evidence on this issue is scarce and does not particularly support the view that liquidity provision by high frequency market makers is particularly more fragile than that of human market makers. Using Nasdaq data, Brogaard *et al.* (2015) find that, on average, HFTs in their sample trade against extreme price movements, whether these movements correspond to permanent (e.g. due to information arrival) or transient price changes. In contrast, non HFTs' orders are positively correlated with the direction of extreme price movements. These preliminary findings suggest that HFTs dampen rather than exacerbate extreme price movements.

The nature of liquidity provision in financial markets might have changed in recent years. In particular, it is possible that liquidity provision is less resilient in case of large shocks. However, again, there might be multiple causes for this evolution and, so far, there is no evidence that HFTs are a direct cause. In fact, banks have cut the amount of capital that they allocate to market making activities in various markets (due to the financial crisis and new regulations such as the Dodd-Frank act in the United States). New players (including HFTs and hedge funds) might replace banks as liquidity providers but these are probably more lightly capitalised and therefore have a smaller risk bearing capacity. This evolution in itself might make liquidity in financial markets more prone to sudden evaporation than in the past.

21 *They define a "breakdown" (resp., breakup) for a stock as a larger than 10% drop (resp., increase) in its price relative to its level at 9:35 a.m. with a reversal of at least 2.5% by 3:55 p.m.*

References

Allen (F.) and Gale (D.) (1994)
"Financial innovation and risk sharing", *MIT Press*.

Back (K.) and Baruch (S.) (2004)
"Information in securities markets: Kyle meets Glosten-Milgrom", *Econometrica*, 72, pp. 433-465.

Baron (M.), Brogaard (J.), Hagströmer (B.) and Kirilenko (A.) (2015)
"Risk and return in high frequency trading", available at: *http://dx.doi.org/10.2139/ssrn.2433118*

Ben David (I.), Franzoni (F.) and Moussawi (R.) (2015)
"Do ETFs increase volatility?", *Working paper*, Ohio State University.

Biais (B.) and Foucault (T.) (2014)
"HFT and market quality", *Bankers, Markets and Investors*, January-February.

Biais (B.), Foucault (T.) and Moinas (S.) (2015)
"Equilibrium fast trading", *Journal of Financial Economics*, 116, pp. 292-313.

Blume (L.), Easley (D.) and O'Hara (M.) (1994)
"Market statistics and technical analysis: the role of volum", *The Journal of finance*, 49(1), pp. 153-181.

Bond (P.), Edmans (A.) and Goldstein (I.) (2012)
"The real effects of financial markets", *Annual Review of Financial Economics* 4, pp. 339-360.

Brogaard (J.), Carrion (A.), Moyaert (T.), Riordan (R.), Shkilko (A.) and Sokolov (K.) (2015)
"High frequency trading and extreme price movements", available at: *https://www.rsm.nl/fileadmin/home/Department_of_Finance__VG5_/LQ2015/Ryan_Riordan.pdf*

Brogaard (J.), Hagströmer (B.), Norden (L.) and Riordan (R.) (2015)
"Trading fast and slow: colocation and market quality", *Review of Financial Studies*, forthcoming.

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

Brunnermeier (M.) and Petersen (L. H.) (2005)
"Predatory trading", *Journal of Finance*, 60, pp. 1825-1864.

Budish (E.), Cramton (P.) and Shim (J.) (2015)
"The high frequency trading arms race: frequent batch auctions as a market design response", *Quartely Journal of Economics*, 130, pp. 1547-1621.

Chaboud (A.), Chiquoine (B.), Hjalmarsson (E.) and Vega (C.) (2014)
"Rise of the machine: Algorithmic trading in the foreign exchange market", *Journal of Finance*, 65, pp. 2045-2084.

Commodity Futures Trading Commission and Securities & Exchange Commission (2010)
"Preliminary findings regarding the market events of May 6, 2010", *Staff Report*, May 18.

Commodity Futures Trading Commission and Securities & Exchange Commission (2010)
"Findings regarding the market events of May 6, 2010", Staff Report, September 30, available at: *https://www.sec.gov/news/studies/2010/marketevents-report.pdf*

Department (US) of the Treasury, Board of Governors of the Federal Reserve System, Federal Reserve Bank of New York, SEC and CFTC (2015)
Joint Staff Report, available at: *https://www.treasury.gov/press-center/press-releases/Documents/Joint_Staff_Report_Treasury_10-15-2015.pdf*

Dugast (J.) (2015)
"Unscheduled news and market dynamics", *Working paper*, Banque de France.

Dugast (J.) and Foucault (T.) (2016)
"Data abundance and asset price informativeness", available at: *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2398904*

ESMA (2014)
"High frequency trading activity in EU equity markets", *Economic Report*, 1, pp. 1-30.

Fama (E.) and Miller (M.) (1972)
"The theory of finance", New York: Holt, Rinehart and Winston.

Financial Times (2015)
"Ex-SEC economist to testify on flash crash", October 22.

Fishel (D.) and Ross (D.) (1991)
"Should the law prohibit 'manipulation'", *Harvard Law Review*, 503.

Fortune (2015)
"How investors are using social medias to make money", December 7.

Foucault (T.), Hombert (J.) and Rosu (I.) (2016)
"News trading and speed", *Journal of Finance*, 71, pp. 335-382.

Foucault (T.), Kozhan (R.) and Tham (W. W.) (2015)
"Toxic arbitrage", available at: *http://dx.doi.org/10.2139/ssrn.2409054*

Foucault (T.), Pagano (M.) and Röell (A.) (2013)
Market liquidity: Theory, evidence and policy, Oxford University Press.

Gao (C.) and Mizrach (B.) (2015)
"Market quality breakdowns in equities markets", available at: *http://dx.doi.org/10.2139/ssrn.2153909*

Grossman (S.) and Stiglitz (J.) (1980)
"On the impossibility of informationally efficient markets", *The American Economic Review*, 70(3), pp. 393-408, June.

Hagströmer (B.) and Norden (L.) (2013)
"The diversity of high-frequency traders", *Journal of Financial Markets*, 16, pp. 741-770.

Hirschey (N.) (2013)
"Do high frequency traders anticipate buying and selling pressure", *Working paper*, London Business School.

Hirshleifer (J.) (1971)
"The private and social value of information and the reward to inventive activity", *American Economic Review*, 61, pp. 561-574.

Huffington Post (2004)
"Twitter causes a flash crash, highlighting market' structural problems", February 23.

Joint Staff Report (2010)
"Findings regarding the market events of May 6, 2010", available at: *https://www.sec.gov/news/studies/2010/marketevents-report.pdf*

Joint Staff Report (2015)
"The US treasury market on October 15, 2014", available at: *https://www.treasury.gov/press-center/press-releases/Documents/Joint_Staff_Report_Treasury_10-15-2015.pdf*

Kumiega (A.), Sterjevski (G.) and van Vliet (B.) (2016)
"Beyond the flash crash: Systemic risk, reliability, and high frequency financial markets", *Working paper*, University of Illinois.

Kyle (A.) and Viswanathan (S.) (2008)
"How to define illegal price manipulation?", *American Economic Review*, 98, pp. 274-279.

Lewis (M.) (2014)
Flash boys – A Wall Street Revolt, W. W. Norton & Company.

Madhavan (A.) (2012)
"Exchange traded funds market structure and the flash crash", available at: *http://dx.doi.org/10.2139/ssrn.1932925*

Security Exchange Commission (SEC) (2010)
"Concept release on equity market structure".

SEC (2013)
"Regulation systems compliance and integrity".

SEC (2014)
Equity market structure literature review, Part II: High frequency trading, staff of the division of trading and markets, pp. 1-37.

Van Kervel (V.) (2015)
"Competition for order flow with fast and slow traders", *Review of Financial Studies*, 28, pp. 2094-2127.

Van Kervel (V.) and Menkveld (A.) (2015)
"High-frequency trading around large institutional orders", *Working paper*, Vrije Universiteit.

Wall Street Journal (2014)
"High-speed stock traders turn to laser beams", February 11.

Weller (B.) (2016)
"Efficient prices at any cost: does algorithmic trading deter information acquisition?", *Working paper*, Northwestern University.

Yang (L.) and Zhu (H.) (2015)
"Back-running: Seeking and hiding fundamental information in order flows", *Working paper*, MIT.

Ye (M.), Yao (C.) and Gai (J.) (2013)
"The externalities of high frequency trading", *Working paper*, University of Illinois.

Zhang (S.) (2012)
"Need for speed: An empirical analysis of hard and soft information in a high frequency world", *Working paper*, University of Manchester.