

Duisenberg
school of
finance



DSF Policy Paper Series

Foresight Driver Review: Electronic Trading and Market Structure

Albert Menkveld

October 2011

DSF Policy Paper, No. 16

Foresight Driver Review: Electronic Trading and Market Structure

Albert Menkveld*

Executive Summary

Securities trading is essentially a search process whereby investors' aim is to identify security transfers where a buyer has a higher utility from owning an asset than the seller does. This utility differential might stem from a private value differential (e.g., hedge values are idiosyncratic to investors) in which case it is referred to as social gains-from-trade. If it stemmed from an informational differential on the security's common value (e.g., the mean of future dividends), then trade price partially reveals the information of the more informed trader. Such information production is useful in allocation scarce resources in an economy.

In a pre-electronic age the search process was particularly costly and a single centralized market could capitalize on a large network externality as this created a high entry barrier for potential rival markets. This changes in the electronic age as search cost has become almost trivial. A price in a potential rival market is easily checked by a computer. The network externality is substantially reduced and the incumbent market becomes contestable. New markets get started and investors benefit from competition in terms of lower fees and innovation, but they also benefit from market heterogeneity as investors differ in their trading needs.

Computers are used to match orders in an electronic market but they are also useful for implementing automated trading strategies, a practice referred to as 'algorithmic trading'. There are two types of algorithms. Agency algorithms implement a long-term position change at the lowest possible cost. Proprietary algorithms trade to lock in a trading profit and positions are relatively short-term. Agents who run such algorithms are commonly known as high-frequency traders (HFT).

Trading cost has shown a general decline in the period of migration to electronic trading, yet it has become more erratic. The bid-ask spread decreased, depth increased, commissions and fees have fallen. This is arguably driven by forces of market competition but potentially also by the ability of computers to reduce adverse selection risk. Also, price quotes appear to become pressured more quickly as traditional market makers with an affirmative obligation to 'lean against the wind' have disappeared. An extreme example is the ash crash on May 6, 2010, when the main U.S. stock index dropped by 8% in afternoon trading and recovered, all in a 15-minute time window. This is particularly worrying as there is preliminary evidence that investors command a larger premium for stocks with frequent 'disaster liquidity' events which raises the cost of capital for issuers.

High-frequency traders could improve the trading outcome for investors. They are good for securities trading to the extent that they are natural suppliers of price quotes. They have invested in software and hardware to parse the public information flow at superhuman speed and refresh their quotes before others can adversely select them. They also benefit trading if they are connected to multiple markets at the same time. They can transfer trade interest across

markets as they become the natural counterparty to a trade in one market if they can offload a position in another market.

High-frequency traders might also negatively affect trading. If HFTs trade aggressively on quickly-processed public information they effectively increase adverse selection on investors' price quotes. This essentially eliminates the ability for investors to earn the bid-ask spread rather than pay it which might make some trading strategies prohibitively expensive (e.g., option replicating strategies). Also, speed might trigger a socially wasteful arms race among high-frequency traders. Finally, electronic markets might be particularly vulnerable to new manipulative or socially destructive trading strategies.

* VU University Amsterdam, Duisenberg school of finance

This review has been commissioned as part of the UK Government's Foresight Project, The Future of Computer Trading in Financial Markets. It was published as DR16, see: <http://bis.ecgroup.net/Publications/Foresight/ComputerTradinginFinancialMarkets.aspx>. I would like to thank Ian Domowitz, Stephen Figlewski, Thierry Foucault, Terrence Hendershott, Charles Jones, Oliver Linton, Emiliano Pagnotta, and Yueshen Zhou for their feedback. All errors are mine.

1. Introduction

The vision presented in this document benefits from many decades of academic research in microstructure and my personal interaction with various leading academics, market participants, and regulators. This vision was commissioned by the British government with an explicit request to "please avoid inserting multi-page bibliographies." It is for this reason that references are kept to a minimum and, as the document is strictly my personal view, they are skewed towards my own work.

I do want to emphasize that most credit should go to those who necessarily go unmentioned below.

Advances in information and communication technology (ICT) have enabled automated trading which, in turn, changes the structure of the securities trading industry. For long, human intermediation was the only feasible way for investors to exchange securities. Broker-dealers traded securities on behalf of their clients in a 'market' that was either the floor of an exchange or an interdealer network. ICT makes human intermediation optional and therefore fundamentally changed the institutional structure of exchanging securities which affects all participants: investors, broker-dealers, markets, and regulators.

The fundamental role of (secondary) securities markets is to produce transactions that either lead to Pareto-improving re-allocations, often referred to as gains-from-trade, or reveal information. Transactions produce gains-from-trade if the buyer has higher private value from owning the asset than the seller does. For example, the asset's payoff might correlate more with the endowment stream of the buyer relative to that of the seller. This way, the private value is produced through a hedge of the owner's income stream. A transaction produces information if at least one side to the trade was motivated by a private signal on the 'fundamental' or common value of the security. This has social value if it leads to a better allocation of scarce resources in an economy, i.e., early information on poor performance of a management team puts pressure on the team to for example exercise more effort, change strategy, or to simply leave.

One important effect of the new electronic environment is that markets have less ability to extract a 'monopoly' rent that capitalizes on a large network externality. In essence, securities trading is a search process where agents' objective is to find a mutually beneficial transaction. An agent's search cost is the effort or cost associated with contacting either an intermediary or a natural counterparty to a trade, for example a call to a broker-dealer. In a human-intermediated market, search costs are substantial and every additional participant to a network substantially reduces the search cost for existing participants. A single centralized market is therefore, in effect, not contestable - it is hard for an entrant to lure investors away (cf. Domowitz and Steil (1999)). A key driver in the transition to a new institutional structure is that electronic communication is cheap. This lowers search cost and markets become contestable. Electronic market access makes checking prices in a competing new market almost trivial. For example, before hitting a bid quote in the incumbent market, an investor might quickly pull-up the bid quote in the entrant market and, if higher, route her sell order to the new market instead. The next step is to automate this process whereby a 'smart router' consolidates all bid quotes across markets into a single screen, the investor decides how much to sell, and the computer routes the order to the markets from which the targeted bid quotes originate.

Search cost reduction has led to a natural process of market entry and competition for order flow. This has both put pressure on the cost of operating a market and led to innovation in the matching process. Such innovation pertains to both the mechanics of matching orders (e.g., limit order book, crossing network, dark pool), matching speed, and fee structure (e.g., rebates for limit orders that add to the limit order book, see Foucault, Kadan and Kandel (2010)). The heterogeneity in market structures in and of itself creates social value if trading needs differ across investors (Stoll (2001)). It also should be noted that an important facilitating factor was the more pro-competitive regulation adopted in the U.S. and in Europe (Reg NMS and MIFID, respectively).

Markets fragment as a result of the new electronic environment but they get partially re-consolidated through algorithmic trading which is commonly defined as “the use of computer algorithms to automatically make certain trading decisions, submit orders, and manage those orders after submission (Hendershott, Jones and Menkveld (2011, p.1).” These algorithms are naturally categorized as agency or proprietary algorithms. Agency algorithms are used by investors to trade into a desired long-term position at minimum transaction cost. The aforementioned smart router is an early manifestation of such agency algorithm. Proprietary algorithms are ‘self-motivated’ in the sense that they buy and sell securities to lock in a trading profit. Agents who run these algorithms are often referred to as high-frequency traders.

Accordingly:

1 In human-intermediated markets search costs are high. A centralized market operator therefore benefits from a sizeable participation externality; it is effectively shielded from competition as the entry barrier is prohibitively high for new markets. Electronic trading reduces search cost, therefore lowers a market's participation externality, reduces the entry barrier for new markets and thus creates real competition among multiple markets.

Figure 1: Overview main idea

This graph depicts the main idea of the manuscript. Panel A conveys the economics of securities trading in the most parsimonious way. Panel B illustrates the essential differences between the traditional institutional structure of the securities trading industry (top) and the modern electronic structure (bottom).

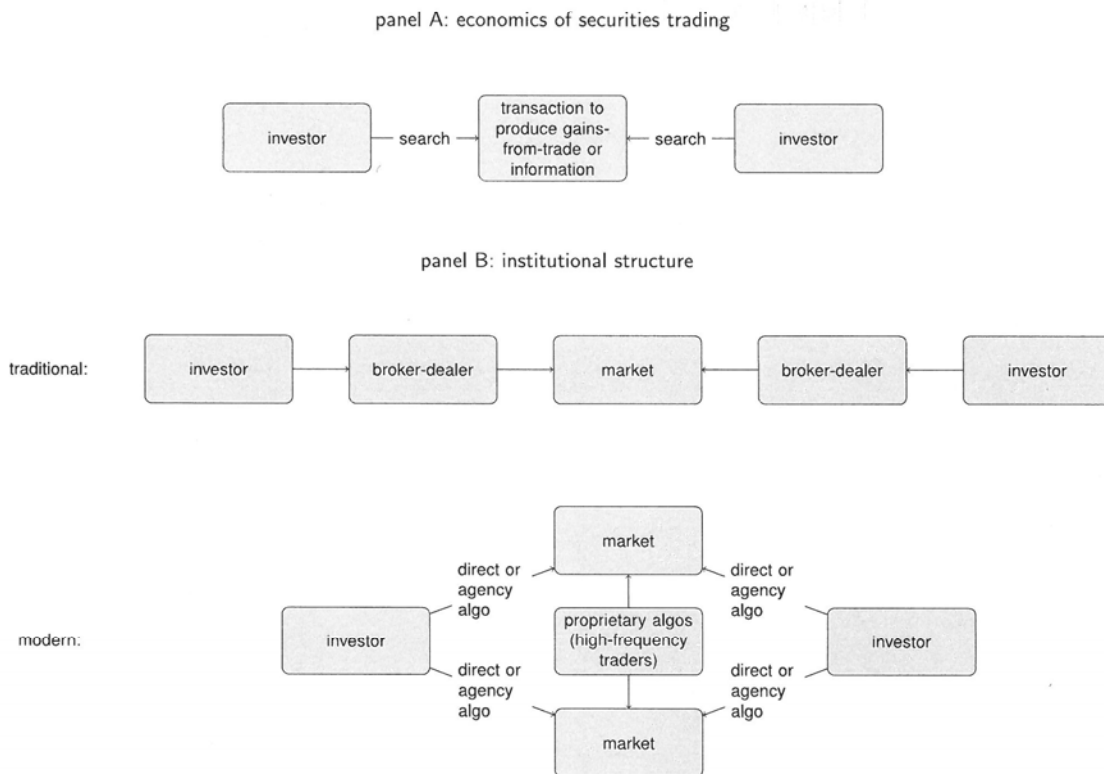


Figure 1 summarizes the vision that has been presented thus far. Panel A graphs the way in which securities trading produces value. Panel B compares and contrasts the traditional and the modern structure of securities trading. This figure serves as an overarching structure to organize the ideas in the remainder of the manuscript.

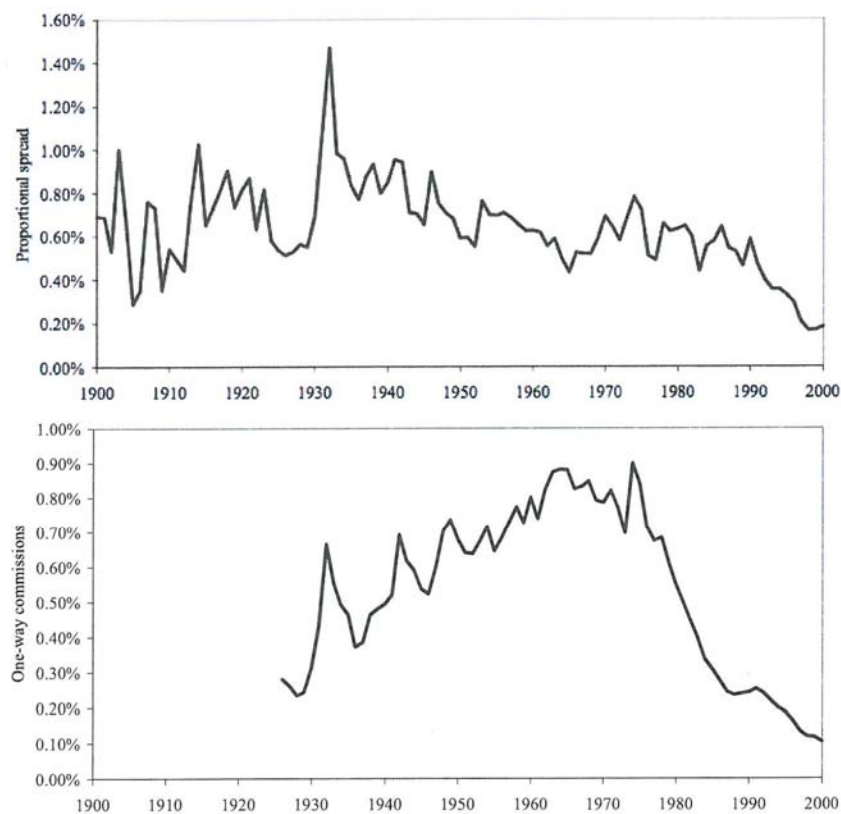
Securities markets with a sufficiently large trader population all seem to transition to electronic trading. Equity markets were among the first to turn electronic and it is for this reason that the evidence in this review is based on these markets. I discuss the most relevant trends and extrapolate from them when discussing the future of securities trading.

2 Trends in 21st century equity trading

The total cost of executing a trade consists of an internal search cost to identify the trade opportunity and a transaction cost that has both an implicit and an explicit part. The implicit cost is the 'price concession' the initiating side of a trade willingly incurs to get a trade done. If a trader buys she pays the ask price; when she sells she gets the bid price. The bid-ask spread is an example of the 'implicit' cost for an investor who initiates a trade and therefore demands liquidity. The explicit cost is the sum of commissions and fees she owes to those whose services are consumed along the way. In a traditional human-intermediated market these are, for example, the commissions that need to be paid to broker-dealers or the exchange fee.

Figure 2: A century of transaction cost

The top graph depicts the relative bid-ask spread (the ask price minus the bid price divided by the average of the two) for NYSE trading in Dow Jones stocks. The bottom graph depicts the relative commissions (commissions divided by the transaction value) paid for trading these stocks.

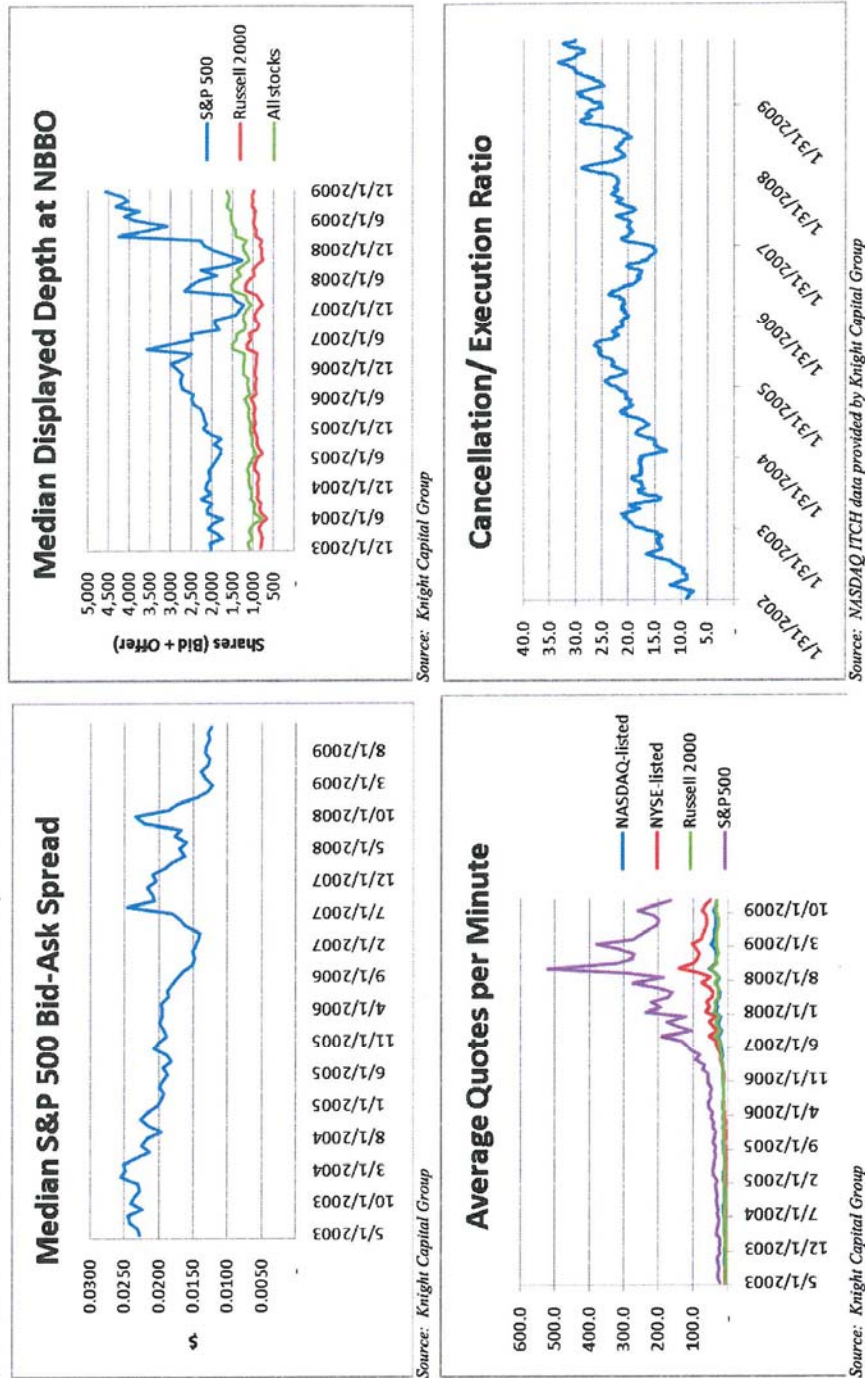


source: Jones (2002)

Jones (2002) shows that in the last two decades of the twentieth century transaction cost declined for Dow Jones stocks. Figure 2 shows that the relative bid-ask spread (i.e., the spread as a percentage of the average of the bid and ask quote) hovered around 0.80% for most of the century but started to decline in the nineties. Commissions, an important explicit cost, rose in the middle of the century from 0.30% to 0.80% but the trend reverses in the late seventies to turn into a gradual decline towards 0.15% at the end of the century, a record low. This downward trend coincides with a gradual transition to electronic trading in this time period.

Figure 3: 21st century equity trading

These graphs summarize trends in 21st century equity trading. The top graphs illustrate the general improvement in liquidity supply (bid-ask spread and depth at the best quotes). The bottom graphs illustrate the rapid increase in quote activity.



source: Angel et al. (2010)

Angel, Harris and Spatt (2010) review equity trading in the first decade of the 21st century and Figure 3 is a selection of the most relevant graphs for the purpose of this manuscript. The top two graphs illustrate that implicit transaction cost continued to decline in the first decade. The top-left graph shows that the bid-ask spread for S&P 500 stocks declined from approximately USD 0.0250 at the start of the decade to USD 0.0125 at the end of it, a reduction of 50%. It does exhibit elevated levels during the 2007-2008 financial crisis but this appears to be entirely event-driven as the bid-ask spread as of 2009 is below pre-crisis levels.

The distance between the best prices available on both sides of the market, the bid and ask price, is a useful yet incomplete measure of implicit transaction cost. An equally relevant measure is how many securities are available for trade at these prices, i.e., a measure of market depth. The top-right graph shows that depth has increased for the S&P 500 stocks from around 2000 shares at the start of the decade to 4500 at the end of it; depth has more than doubled. The lower depth levels in 2007-2008 are consistent with the higher bid-ask spread in this period, i.e., they both indicate a reduced liquidity supply. Coincidentally, most of the transition to electronic trading in U.S. equity markets was concentrated in this decade. The most prolific example is the NYSE transition from mostly a floor market to an almost no-floor market in the course of the decade.

Accordingly:

2 There is evidence that transaction cost declined throughout the period of transition to electronic trading.

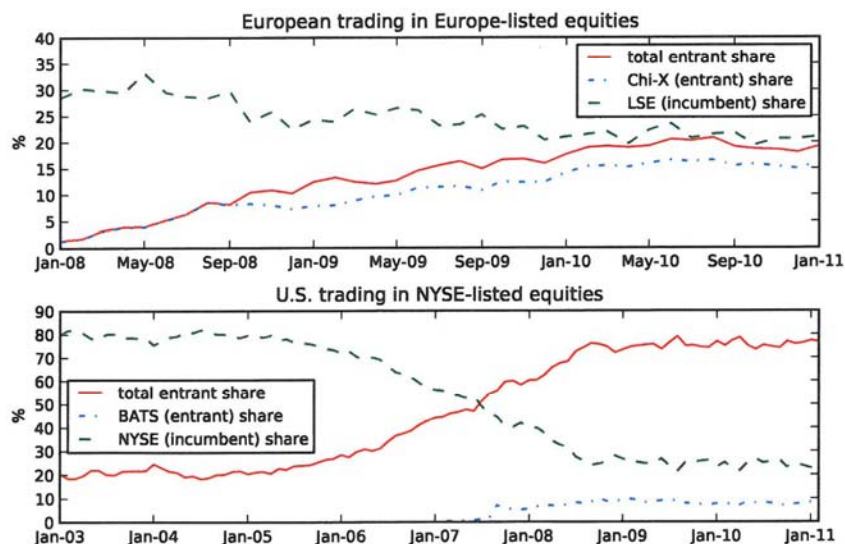
The other salient trend in the first decade of equity trading is a sharp increase in quote activity. The bottom graphs in Figure 3 illustrate this trend. The bottom-left graph shows that the average quotes per minute per stock has increased from 20 to 200 with even higher levels reached during the 2007-2008 financial crisis. Admittedly, this increase might just reflect an increase in the number of trades as a result of a general transaction cost decline. But, the bottom-right graph shows that this cannot be the full explanation; the cancel-to-trade ratio has increased from 10 to 30 in this period. The lower increase in this ratio relative to the quote activity intensity shows that the activity increase is both due to more trades and more quote cancellations in between trades.

Accordingly:

3 Market activity has increased sharply in the first decade. The number of trades has risen as well as the number of quotes per trade.

Figure 4: Market fragmentation

These graphs illustrate the entry of new markets and the subsequent fragmentation of trading. The top graph presents the total entrant share of EUR volume in Europe-listed stocks as well as the largest entrant (Chi-X) share and the largest incumbent (London Stock Exchange) share. The bottom graph presents the total entrant share of share volume in NYSE-listed stocks as well as a large entrant (BATS) share and the incumbent (NYSE) share.



source: Barclays Capital Equity Research and Federation of European Securities Exchanges

Figure 4 illustrates that both in Europe and in the U.S. new entrant markets have captured significant market share towards the end of the decade. In Europe, the new entrants are Chi-X, BATS, and Turquoise who collectively grew from a zero market share at the start of 2007 to a 20% share in January 2011. In the U.S., the most striking example is that NYSE market share in NYSE-listed stocks fell from 80% in January 2003 to about 25% in January 2011. This lost market share appears relatively equally distributed among various competing platforms. For example, one of the larger ones is BATS which in January 2011 had a 10% market share.

Accordingly:

4 New markets obtained significant market share in the second half of the first decade both in the U.S. and in Europe.

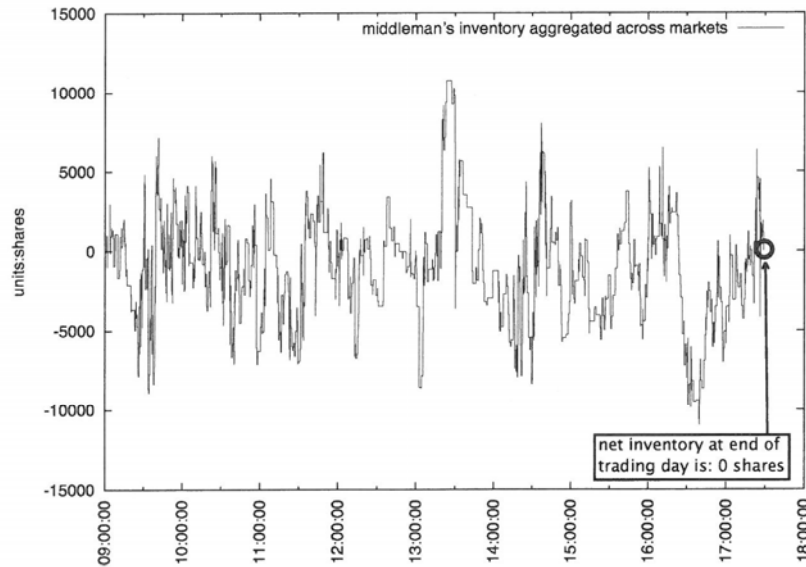
At about the time that the new platforms made significant inroads into securities trading, proprietary algorithms operated by high-frequency traders (HFT) started to flourish. Their participation in trades is often estimated to be higher than 50%. For example, Brogaard (2010) documents a participation rate of 73.7% for a 2010 sample of NASDAQ stocks. Larry Tabb, chief executive of Tabb Group, a consultancy, said high-frequency trading accounted for 54% of U.S. equity trading and 35% of European equity trading¹. HFTs or proprietary algorithms, I will use these labels interchangeably, are characterized as automated trading strategies implemented on lightning-fast computers that operate at the shortest possible distance from the matching engine of a market. The objective of proprietary algorithms is to lock in a trading profit. Agency algorithms, on the other hand, are used to establish a

¹ See "High-frequency boom time hits slowdown," Financial Times, April 12, 2011

long-term position change at the lowest possible transaction cost (by, for example, splitting a large order optimally across markets and across time). HFT typically operate on intraday frequencies; they start and end the day on a zero position (see SEC (2010a) for further details on HFT characterization).

Figure 5: One day in the life of a high-frequency trader

This graph illustrates the nature of high-frequency trading. It depicts the position of one broker ID in ING stock on January 30, 2008. The position is the aggregate net position across both the incumbent market, Euronext, and an entrant market, Chi-X.



source: Jovanovic and Menkveld (2010)

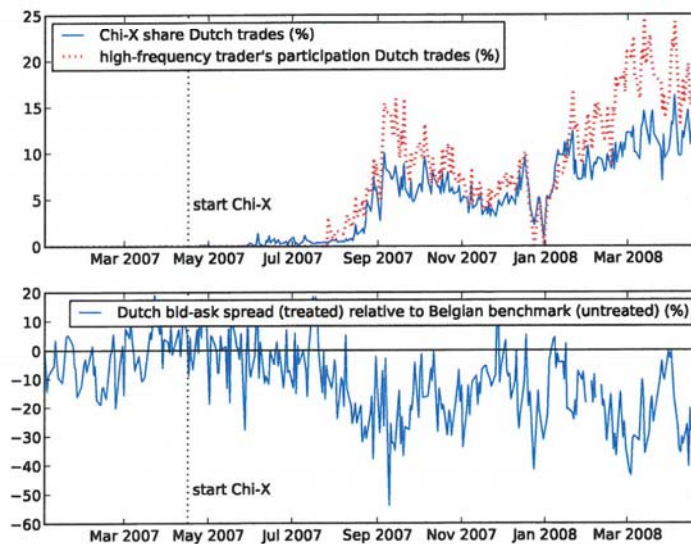
Figure 5 illustrates the operation of an HFT algorithm on a large Dutch index stock. It trades the stock in two markets, the incumbent market Euronext and the entrant market Chi-X. The HFT starts at a zero position, trades thousands of shares throughout the day, and ends at a zero position. It is not only plain-vanilla arbitrage across the two platforms as that involves buying and *simultaneously* selling in case the bid in one market is higher than the ask in the other. In that case the HFT position should be zero throughout the day whereas the graph shows that its position ranges between short 10,000 shares and long 10,000 shares. A nonzero position seems to last up to 30 minutes.

Accordingly:

5 High-frequency traders' (i.e., proprietary algorithms) participation in trades was small in the middle of the decade but, reportedly, grew to 30-70% by the end of the decade.

Figure 6: **High-frequency trader activity, entrant market share, and the bid-ask spread**

This figure depicts the evolution of trading in Dutch index stocks in the incumbent market Euronext and the entrant market Chi-X. The top graph plots the market share of Chi-X along with the participation rate (in both markets) of a new and large high-frequency trader. The bottom graph plots the average bid-ask spread of 'treated' Dutch stocks relative to 'untreated' Belgian benchmark stocks.



source: Menkveld (2011)

Figure 6 suggests that the success of the new platform Chi-X (observation 4) and the surge in high-frequency trading (observation 5) are related. The figure depicts the entrant market share in Dutch stocks as of the day of its inception, April 16, 2007. Interestingly, market share hovered around 1% in the early months of operation until, suddenly, in August it grew to a double-digit share. The same graph plots the overall market share of a new broker ID that operates in both Chi-X and Euronext.

It first appeared in the sample in August of that year. By the end of the month it participated in 10-15% of all trades and about half of its trades materialized in Chi-X and the other half in Euronext (see Menkveld (2011)). This broker ID has all features of an HFT and it is in fact the same as the one used to illustrate a high-frequency trader's operation in Figure 5.

The high-frequency trader, whose arrival appears intimately linked to the success of Chi-X, can be characterized as an electronic market maker. In most of its trades the HFT is on the passive side, i.e., hers is the limit order in the book that finds execution on an incoming liquidity-demanding order. A decomposition of its profit shows that the earnings all arrive from the passive orders (through rebates and earning the spread) as, on average, the HFT loses on the inventory positions it carries through time. Further details on the nature, profitability, and capital intensity of the HFT operation can be found in Menkveld (2011).

This evidence on the start of Chi-X supports the vision that was illustrated in Figure 1. A new market successfully enters because high-frequency traders can offer liquidity in the new system (a bid or ask

quote) and, once hit, it will use both the entrant and the incumbent market to offload the position. In this example, the entrant market was superior to the incumbent market as its fees were lower and it ran a faster system. These features are particularly important for a high-frequency trader which, by the nature of its operation, pays the fee (or earns the rebate) frequently and cares about speed in order to change its price quotes quickly on incoming information.

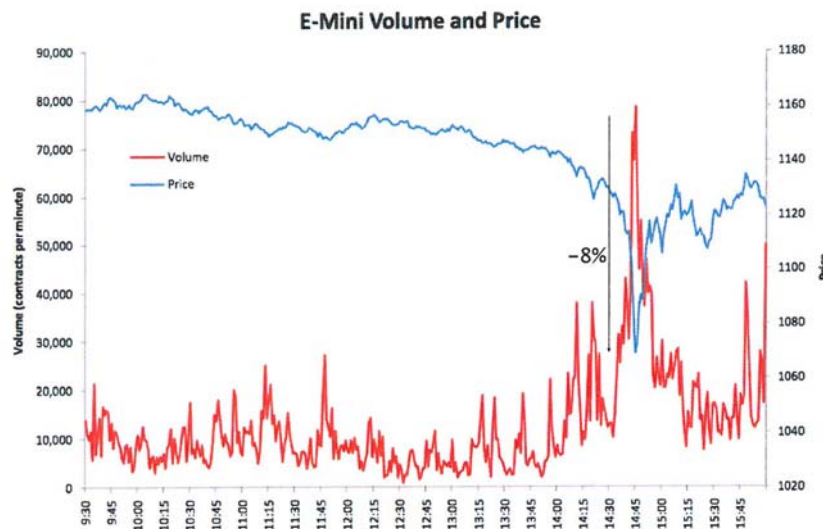
Accordingly:

6 Entrant markets benefit from the presence of high-frequency traders and high-frequency traders thrive on the competition between markets; there appears to be a symbiotic relationship between HFT and new market entry.

The competition for order flow that started with Chi-X entry had an immediate effect on both market quality and trading fees for investors at large. Euronext responded to the fast new system by significantly upgrading the speed on its own system. It also reduced its fees to shrink the gap with the fees charged by the new market. The strongest effect, however, was the reduction in clearing fees. With the entry of Chi-X a new clearing house entered, EMCF, and its aggressive fee structure triggered a fee war with Euronext's LCH-Clearnet. Within a year fees were 50% lower. This constituted a substantial cost saving to market participants as clearing fees are the same order of magnitude as the fees charged by market operators. In addition to these explicit trading cost reductions, the bottom graph of Figure 6 shows that also implicit cost declined as the bid-ask spread was reduced by about 30%.

Figure 7: Flash crash on May 6, 2010

This figure illustrates what has come to be known as the flash crash, the sudden and dramatic decline in the value of the S&P 500. It depicts both price and volume for the E-Mini futures contract that tracks the S&P 500.



source: SEC (2010b)

The migration to electronic trading coincided with a gradual decline in average transaction cost (cf. observation 2) but it also coincided with the occasional disappearance of liquidity supply. The most dramatic example is the ash crash of May 6, 2010, when a large fundamental seller, reportedly, demanded substantial liquidity by trying to sell \$4.1 billion worth of the E-mini futures contract on the S&P 500. It initiated the sell order at 2:32 p.m. through an agency algorithm that was set to work the order aggressively as it, for example, was not restricted by a price limit. Figure 7 illustrates the E-Mini

price path and shows that by 2:45 p.m. the price decline relative to the start of the day was 8% and most of it was after the initiation of the sell order. Not only were E-Mini prices affected, the sell order also caused similar price declines in related index-based securities, and even larger declines in some of the securities in the index (see SEC (2010b)).

A natural reason for, at times, shallow liquidity supply is that machines avoid large positions that might take long to unwind. Such positions are inherently costly as they carry price risk but also, to the extent that they are carried overnight, they tie up capital due to end-of-day margin requirements. A further reason for erratic supply is that machines might withdraw from the market if they observe price paths that they do not 'understand', e.g., a path that does not match any historical path so that the machine does not have an informational (predictive) advantage. In a traditional market, the exchange often designated a market maker who had the affirmative obligation to 'lean against the wind' in particular at times when others withdraw from liquidity supply. HFTs are under no such obligations.

A recent study shows that investors increasingly care about tail risk in liquidity supply as evidenced by higher required returns. Menkveld and Wang (2011) measure such 'tail risk' by identifying liquidity leaks, or short, liquileaks, through estimation of a regime-switching model. A liquileak is defined as the event that an investor funds the stock in a very poor liquidity state for more than a week. In the cross-section, a one standard deviation increase in liquileak probability commands an additional annual premium of 1.33% based on a 1963-2008 sample of 2147 U.S. equities. More importantly, the liquileak premium increased over time.

Accordingly:

7 Transaction cost declined on average (cf. observation 2) yet appears to feature more frequent and more extreme spikes when migrating to electronic trading. The 2010 May 6 ash crash is a dramatic example. A recent study shows that tail risk in liquidity supply seems to increasingly matter to investors. They command a premium for it which raises the cost of capital for the issuing firm.

3 Looking forward: the challenges of electronic trading

In summary, the migration to electronic trading coincides with:

1. Real competition for order flow among a variety of markets as a lower participation externality reduces barriers to entry.
2. A potentially more indirect way in which investors' trade interests are matched. Investors fund one another by either algorithmically splitting their order across markets and through time or through the use of their preferred market in which case middlemen (i.e., high-frequency traders) ensure that they are matched across markets. In the latter case, an investor's order naturally funds the HFT on the other side of the trade as HFTs have the technology to access other markets to search for an offsetting trade. In effect the HFT 'arbitrages' the various markets.
3. A seemingly more erratic liquidity supply. Bid or ask quotes become pressured more quickly as the new intermediaries, high-frequency traders, do not have affirmative obligations to supply liquidity. At the daily frequency there is evidence that investors increasingly care about the probability of 'disaster liquidity' as the premium for it has risen over time.

In the remainder of the manuscript this vision is expanded on to identify future challenges to a well-functioning financial market. Such market either produces security transfers to realize gains-from-trade or it produces socially useful information.

Does the presence of high-frequency traders affect investors' welfare at a more fundamental level than its direct effect of producing competition between markets? Arguably an important advantage of machines over humans is their ability to process vast amounts of information at superhuman speed.

An important information source is price quotes in, for example, index futures, same industry stocks, oil, or fixed income. Machine-powered HFT strategies are therefore likely to interact with the information asymmetry friction in securities trading. This friction and its consequence is best illustrated with a simple example (the formal treatment is in Jovanovic and Menkveld (2010)).

Let us consider a market with investor arrivals that are spaced out in time (cf. Grossman and Miller (1988)) and, for simplicity, let us consider two investors: an early-arriving investor who owns an asset and puts it up for sale and a late-arriving investor who considers buying it. The seller posts an ask price and leaves the market. The buyer arrives and will only buy the asset if her valuation for it is higher than the ask price. The informational friction arises when in between their arrivals some information comes out on the (common value of) the security. Naturally, only the potential buyer, the late investor, is aware of this information. If the information was good, the buyer is more inclined to take the asset as the (unaltered) price now appears relative low. If it was bad, the seller is more likely to be stuck with the asset. The seller is therefore 'adversely selected' by the buyer. Ahead of time, the seller is aware that this might happen and therefore raises her price relative to a non-information benchmark to protect herself. As a result, trade is less likely and gains-from-trade are therefore less likely to be realized. The market therefore functions suboptimally in the case of this informational friction.

The presence of competitive high-frequency traders with their ability to track the interarrival information can restore some of the lost gains-from-trade. The mechanism responsible for such effect is that high-frequency traders bid for the asset at the time the seller is in the market. Price competition (Bertrand) ensures that the HFT bid price is such that ex-post the HFT earns zero rent. The seller can still post an ask instead of accepting the HFT bid which guarantees that HFT presence does not deteriorate her position. The benefit of passing the asset off to the HFT is that machines have the ability to update their ask quote quickly on incoming information in which case *their* ask quote is not adversely selected by the buyer. This could explain observation 3 which states that both the number of trades and the number of quotes per trade have increased. The number of trades increases as it becomes more likely that the seller transfers the asset to the buyer with a potential additional trade if it reaches the buyer via the HFT. The number of quotes per trade increases as the HFT updates its ask quote on incoming information.

The presence of high-frequency traders might, however, lower the gains-from-trade if the late investor, the buyer, was not aware of any interarrival information. In this case, the market did not suffer an informational friction to begin with. Entry of HFTs now creates adverse selection cost for the seller as the machine will snap up the asset on positive news. The seller rationally protects herself by raising the offer price which, in effect, reduces the likelihood of trade. There is therefore less scope to realize gains-from-trade.

One might argue that in this case HFT entry reduces realized gains-from-trade, but still benefits the market in the sense that information gets into prices quicker, i.e., the market has become more efficient. The concern is whether this type of informational efficiency is socially productive in particular when such information would have become available to investors at lower frequencies. For example, an actively traded stock produces a couple of transactions per minute. The interarrival time of investors is therefore in the order of seconds. In such time frame there arguably will be price changes in the index futures market of which a human investor is unlikely to be aware. HFT will use such information to 'pick off' human quotes. But, such information would have become available to investors at no cost at lower frequencies, perhaps hours or at least at the daily frequency. The point is that it is not the type of information that analysts push into stock prices after interviewing a company's new management team, learn about their quality, and recommend buy, sell, or hold. Such information would not have been produced in the absence of analysts (or it would become available at such low frequency, e.g., after quarters of disappointing sales, that the damage has already been done).

Accordingly:

8 The presence of high-frequency traders in electronic markets improves welfare when they act as liquidity suppliers and thereby reduce the informational friction that exists between nonsynchronously arriving investors. It, however, reduces welfare when HFTs pick off investors' quotes at superhuman speed on information that would have been revealed to investors at a somewhat lower frequency.

In addition to superior information processing capacity, a high-frequency trader also pays a market to co-locate its trading robot so as to operate as close as possible to the matching engine. In other words, the HFT pays to have a 'fast door' to the market. Such privilege might or might not be good for the market as a whole. To gauge its effect it is useful to condition on one of the two channels through which HFT affects investor welfare as identified in observation 8. If HFT uses the fast door to adjust its quotes on incoming information to avoid adverse selection, it is good for investor welfare. If, however, HFT uses it to aggressively pick-off other investors' quotes that are slow to change as investors travel through the main door, it is likely to destroy welfare. It does so when co-location for aggressive orders produces a social loss | due to foregone trades as some investors might leave the market if the picking-off cost is higher than the expected gain from trade | is larger than the social gain generated by HFTs who are more likely to take on larger positions when supplying liquidity if they know that they can unwind them more quickly (through aggressive orders).

Among high-frequency traders, speed might become the source of a socially wasteful arms race. If one HFT invests in hardware or software to be microseconds (one microsecond is one millionth of a second) faster to trade on a signal before a rival HFT can, nothing is necessarily gained in terms of social value. True, the information is revealed quicker but at a time scale that is unlikely to yield improved allocations. More importantly, a single HFT does not internalize the relative speed disadvantage it creates for others, i.e., HFT technology investment is the source of a negative externality. As a result HFTs overinvest in machines relative to a social optimum (on the assumption that there was no IT underinvestment in the absence of HFT). For society at large, however, improved computing technology might have social value if it is used to, for example, produce better medical equipment or improved gaming experience. :-).

Accordingly:

9 High-frequency traders overinvest in technology relative to a social optimum if the main motivation is to be ahead of rival HFTs when trading on a publicly observed signal.

The speed of electronic trading requires real-time monitoring of positions to ensure market integrity. For example, the environment makes the well-known 'doubling strategy' increasingly feasible. In roulette, such sure-win strategy is: bet on black, and keep doubling your bets until black comes out (see Harrison and Kreps (1979)). The strategy might require access to (infinitely) large credit and (infinitely) many trading rounds in case the player is unlucky in the early rounds. These conditions are increasingly met when markets operate at high frequency and risk controls operate at considerably lower frequency (e.g., some clearing houses' base their margin calls on end-of-day positions with relatively ad-hoc intraday monitoring). A high-frequency trader might implement such doubling strategy by, for example, at the start of the day betting that sometime in the day the price will be higher than the opening price, i.e., bet on 'black'. It starts the day with buying an initial position and keeps doubling it on falling prices until the market 'hits black'. It unwinds the position at that price and pockets a trading profit. Admittedly, such strategy requires negligible market impact when trading but for a thick enough market and a smart execution this might work. The strategy earns a 'risk-free' profit until, one day, doubling might be required so often and time might run out before the market hits black. In this case, the HFT will not have sufficient funds to meet the end-of-day margin call and the central clearing house (and therefore all other participants or the government) will have to take the loss.

Accordingly:

10 *High frequency electronic markets with low frequency system-wide monitoring increasingly allows for destructive doubling strategies.*

References

Angel, J. J., Harris, L. E. and Spatt, C. S. (2010), Equity trading in the 21st century, Manuscript, Carnegie Mellon.

Brogaard, J. A. (2010), High-frequency trading and its impact on market quality, Manuscript, Kellogg School of Management.

Domowitz, I. and Steil, B. (1999), Automation, trading costs, and the structure of the trading services industry, Manuscript, Brooking-Wharton Papers on Financial Services.

Foucault, T., Kadan, O. and Kandel, E. (2010), Liquidity cycles and make/take fees in electronic markets, Manuscript, HEC Paris.

Grossman, S. J. and Miller, M. H. (1988), 'Liquidity and market structure', *Journal of Finance* **43**, 617-633.

Harrison, J. M. and Kreps, D. M. (1979), 'Martingales and arbitrage in multiperiod securities markets', *Journal of Economic Theory* **20**, 381-408.

Hendershott, T., Jones, C. M. and Menkveld, A. J. (2011), 'Does algorithmic trading improve liquidity?', *Journal of Finance* **66**, 1-33.

Jones, C. M. (2002), A century of stock market liquidity and trading costs, Manuscript, GSB Columbia University.

Jovanovic, B. and Menkveld, A. J. (2010), Middlemen in limit-order markets, Manuscript, VU University Amsterdam.

Menkveld, A. J. (2011), High frequency trading and the new-market makers, Manuscript, VU University Amsterdam.

Menkveld, A. J. and Wang, T. (2011), How do designated market makers create value for small-caps?, Manuscript, VU University Amsterdam.

SEC (2010a), 'Concept release on equity market structure'. Release No. 34-61358; File No. S7-02-10.

SEC (2010b), 'Findings regarding the market events of may 6, 2010'.

Stoll, H. R. (2001), 'Market fragmentation', *Financial Analysts Journal* **57(4)**, 16-21.