*Quantitative Finance*, 2020
Vol. 20, No. 11, 1779–1794, https://doi.org/10.1080/14697688.2020.1748700

# The information content of high-frequency traders aggressive orders: recent evidence†

PAMELA SALIBA*

École Polytechnique, CMAP and Autorité des Marchés Financiers, Paris, France

(*Received 1 April 2019; accepted 17 February 2020; published online 13 May 2020*)

This empirical study uses a unique recent data set provided by the French regulator 'Autorité des Marchés Financiers' and gives some evidence concerning the impact of aggressive orders on the price formation process and the information content of these orders according to the different order flow categories (high-frequency traders, agency participants and proprietary participants). We find that both the instantaneous and the transient price impact of aggressive orders consuming exactly the quantity present at the best limit is higher than that of the ones consuming less than the quantity present at the best limit. Furthermore, the price impact is an increasing function with respect to the consumed share in percentage. We show that these price impact disparities are sustainable over time: both price impacts are permanent. In contrast to previous literature, we find that aggressive orders of HFTs carry more information about the short-term behaviour of the price than the ones of agency and proprietary members. This new finding may be an indicator of the evolution of high frequency traders activity over the years.

*Keywords*: High-frequency trading; Aggressive orders; Price impact; Price formation; Asymmetric information; Price profile; Mean reversion; Trend following; Market microstructure

## 1. Introduction

Since the emergence of high-frequency traders (HFTs), a lot of academic research and regulatory discussions have investigated their behaviour and their impact on the markets. It is traditionally considered that the main activity of HFTs is market making, to the extent that HFTs are described as the new market makers in Menkveld (2013). Market makers are defined as market participants who provide liquidity to the market by posting simultaneously limit orders on both sides of the electronic Limit Order Book (LOB), see Brogaard *et al.* (2014) and Jones (2013). This is why HFTs limit orders were widely studied in the literature, see, for example, Cardaliaguet and Lehalle (2017), Chordia *et al.* (2001), Jovanovic and Menkveld (2016), Lehalle and Mounjid (2017) and Megarbane *et al.* (2017). However, HFTs do send aggressive orders, even when they carry out market making strategies, see for instance Jones (2013) and Megarbane *et al.* (2017). In the literature, it has been found that HFTs aggressive orders are not significantly more informed than those of other participants, see Biais *et al.* (2016) where the authors use data going back to 2010, see also Brogaard *et al.* (2014) for related results on date going back to 2008–2009. In this study, we investigate whether this conclusion remains true nowadays by analysing more recent data, from 2017. We furthermore aim at understanding how aggressive orders impact the price.

We first study the price impact of aggregated aggressive orders. Note that there are mainly two types of price impact that are studied in literature: the one of aggregated orders, and the one of single orders, see Bouchaud *et al.* (2009). Aggregation means that the market impact is conditioned on a given number of trades or to a given interval of time, while a single order means that a large order is split into smaller pieces before being sent to the market. The price impact of meta-orders is in general proved to follow a power law, see for example Almgren *et al.* (2005), Moro *et al.* (2009) and Tóth *et al.* (2011) while the results of aggregated orders are more variable from a study to another, and depend mainly on the studied time scale. Dependence on time-scale is discussed in this paper; we study the price impact of aggregated orders on different time scales: from one microsecond till 17 min after the aggressive order. Furthermore, we aim at looking at how the price impact of aggregated aggressive orders varies according to the fraction of consumed liquidity against that present at the best limit. This is why we split aggressive

*Email: pamela.saliba@polytechnique.edu
†This paper originates in the technical note (Saliba 2019a) from 'Autorité des Marchés Financiers'.

orders in three groups: the ones consuming less than the liquidity available at the first limit (partial aggressive orders), the ones consuming exactly the liquidity available at the first limit (exact aggressive orders), and the ones consuming more than the liquidity available at the first limit (N-limit aggressive orders). We evaluate the price impact (in magnitude and durability) conditionally on these groups, by using the notion of price profile. The price profile is the evolution of the price in time around a specific market event. It is in general used to quantify the adverse selection like in Lehalle and Mounjid (2017) or to estimate the information content of orders like in Biais *et al.* (2016). We notice that the repartition of HFTs aggressive orders in these groups is not the same as the repartition of other order flow categories: they send a significant larger proportion of exact aggressive orders than the rest of the market. For partial aggressive orders, we show that the price impact is increasing according to the consumed share in percentage: it depends on the traded volume and the quantity present at the best limit. Our first main new empirical finding is that the disparities in price impact between the three groups of aggressive orders are sustainable over time.

Second, we investigate whether HFTs have an informational advantage compared to the rest of the market. To assess their informational advantage, we compute the potential profits (the potential profit being a proxy of the Pnl) of the different order flow categories (HFTs, agency participants and proprietary participants) over different time horizons, and compare those of HFTs to the rest of the market. We compute the potential profit on different time horizons in order to show how the information content could vary with time: some members are more informed on a short-term horizon, while others are more informed on a long-term one. The potential profit is computed using the price profile: we look at the price variation after the aggressive order compared to the price obtained by the aggressive order. In contrast to Biais *et al.* (2016), we find that HFTs are the most profitable agents on a short-term horizon. The divergence of these results may be an indicator of the evolution of HFTs activity in the market over time. Despite these different results, our third main finding is in line with (Biais *et al.* 2016): HFTs typically buy after price decreases and sell after price increases, while agency members sell after price increases and buy after price decreases. Additionally, we find that partial aggressive orders are more discriminating than exact ones in terms of potential profit disparities between the different order flow categories. Furthermore, we show how dissociating flows of a given market member according to their connectivity channels allows us to exhibit different significant 'sub-behaviours'.

Third, we display how aggressive orders can be used in order to classify market participants as HFTs or non-HFTs. This is of particular interest from a regulatory viewpoint. We document other statistical features, such as the autocorrelation of aggressive orders or specific price patterns related to the behaviour of certain market participants.

This paper is organised as follows. In Section 2, we present our data, describe the different order flow categories and explain our methodology to identify HFTs. We introduce in Section 3 the notion of price profile used to quantify price impacts and potential profits. In Section 4, we distinguish between three different aggressive order groups, and display the price impact of each of them. We then focus on analysing one specific group: the partial aggressive orders according to the consumed share. We measure in Section 5 the potential profit of each order flow categories. In Section 6, we shed light on some statistical features and market participants classification criteria. We even show that we can achieve a more granular and relevant classification using connectivity channels. Finally, Section 7 summarises our results.

## 2. Data description and HFT identification

### 2.1. Data description

We recall that our data are provided by the French regulator 'Autorité des Marchés Financiers' (AMF). This analysis is conducted on the CAC 40 stocks traded on Euronext Paris over a three-month period: from September 2017 to November 2017, during which the volatility on the CAC 40 was stable and reached historically low levels (see Figure 1). Furthermore, this studied period is neither disrupted by end of year trading effects nor by MIFID II.†

In Figure 1, we plot the Vstoxx from 2013 till 2017. The Vstoxx is the 'European VIX'. It measures implied volatility of near term EuroStoxx 50 options, which are traded on the Eurex exchange.

This study focuses solely on the analysis of strategies on Euronext, and does not consider other exchange platforms. Over the analysed period, we use both trade data and LOB data to describe the dynamics of the LOB accurately before and after each aggressive order. The whole data set contains approximately 8 millions aggressive orders and 423 millions events (an event can be an order insertion, an order cancellation, an order modification or a transaction). Note that we do not use market data corresponding to the initial and final twenty minutes of the trading session, as these periods usually have specific features due to the opening/closing auction phases.

### 2.2. The different order flow categories

The data we have gives us access to the order flow category to which each aggressive order belongs. We have four different categories of order flows:

- Agency flows: it corresponds to aggressive orders triggered by market participants acting for the account of their clients.
- Proprietary flows: it corresponds to aggressive orders triggered by market participants acting for their own account.
- Supplemental Liquidity Provider (SLP) flows: it corresponds to aggressive orders declared as part of the SLP programme to which the market participant initiating the order must belong to. The

† MIFID II is a legislative framework instituted by the European Union to regulate financial markets, that entered into force in January 2018.

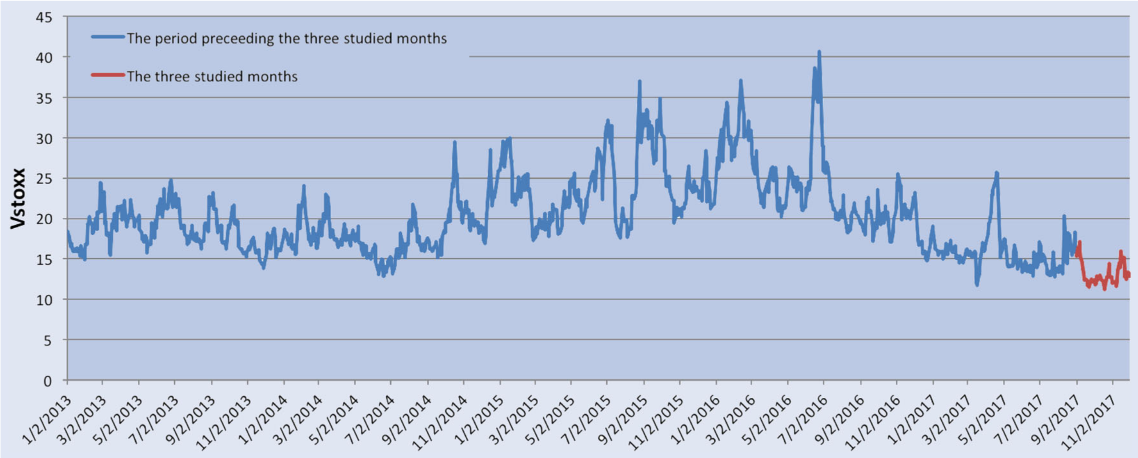


Figure 1. The implied volatility (Vstoxx) during the studied period varies little. The three months under study show the lowest implied volatility since 2013.

SLP programme imposes a market making activity on programme members, including order book presence time at competitive prices. In return, they get favourable pricing and rebates in the form of a maker-taker fees model directly comparable to those of the major competing platforms, see Euronext (2016).

- Retail Member Organisation (RMO) flows: RMO members are part of a programme offered by Euronext: the Retail Matching Facility, specialised in providing liquidity for retail participants through Retail Liquidity Provider (RLP) members. The role of RLP members is to provide liquidity to RMOs by posting buy and sell limit orders. The RMO members are eligible to trade with all market participants, while RLPs can trade only against RMO orders. In the following, the RLP are not considered since they almost never send aggressive orders.

Note that an institution can have one single or multiple member codes to access the market. In addition, using the same member code, an institution can send orders belonging to different order flow categories. For instance, a same institution with the same member code can have at the same time agency and proprietary activities. The data we have provide us with the name of the institution and the member code issuing each order. Additionally, each order is labelled by the order flow category to which it belongs (agency, proprietary, SLP or RMO).

### 2.3. *HFT identification*

High-frequency trading is a subset of algorithmic trading (MIFID II states that algorithmic trading means trading in financial instruments where a computer algorithm automatically determines individual parameters of orders such as whether to initiate the order, the timing, price or quantity of the order or how to manage the order after its submission, with limited or no human intervention) for which minimising latency is a crucial element for performance. HFTs use co-location and proximity services to minimise latency. Most of them submit large numbers of orders that are cancelled relatively shortly after submission, trade large volumes, consistently maintain a low inventory level by holding positions for very short time and turning them over rapidly, see for example Brogaard *et al.* (2014) and Megarbane *et al.* (2017).

In this work, an order is identified as belonging to the high-frequency traders order flow category if it is labelled as SLP. Indeed, HFTs are now essentially the only market participants that are able to play the role of market makers on liquid stocks, see Brogaard *et al.* (2014) and Jones (2013). This is because they are supposingly able to maintain a strong presence at best price limits and operate efficient inventory management in an increasingly fast-moving and fragmented market. Indeed, HFTs can use speed to enhance risk control by avoiding adverse selection, see Jovanovic and Menkveld (2016), improving inventory management, see Ait-Sahalia and Sağlam (2017) and trading on short-lived information, see Foucault *et al.* (2016). Moreover, according to the classification based on the lifetime of cancelled orders (described in details in Section 6.2) which constitutes one of the criteria used by AMF to identify HFTs, all SLP members are either classified as HFTs or mixed HFTs (investment banks with high-frequency trading activity). The analysis of aggressive orders presented in Section 5.3 is thereafter generalised in Appendix to all the HFTs based on this classification, and we show that the results are similar to those obtained when considering only SLPs.

## 3. Quantification of the price impact and the informational advantage

Let us consider a buy (resp. sell) aggressive order occurring at time $t$. To estimate its information content at time $t+h$, we compute the potential profit denoted by $PP_{t+h}$ that a market participant can realise if he succeeds to unwind his position passively:

$$PP_{t+h} = \frac{BP_{t+h} - P_t}{S} * sign_t,$$

where $BP_{t+h}$ is the best ask (resp. bid) at time $t+h$, $P_t$ the price per share obtained by the aggressive order, $sign_t$ takes the value 1 (resp. $-1$) if it is a buy (resp. sell) aggressive order. The quantity $S$ is the average spread of the asset. It is computed by averaging among all the events in the data set happening between 9:20 and 17:10, and weighted by time. Finally, $h$ varies between $-17$ min to 17 min. Note that when computed at time horizons before the aggressive order, this measure does not reflect the potential profit. However, it will help us to understand the strategy followed by the market participants (mean reverting or trend following), see Section 6.4. On the other hand, it will also allow us to determine whether the aggressive order has been sent at a relevant time or not. At a given point in time before the aggressive order, a negative value indicates that the participant could have obtained a better price (at least for one security). On the contrary, a positive value indicates that the participant has intervened at a convenient moment: if the aggressive order had taken place earlier, the price would have been worse.

Note that it would have been possible to quantify the potential profit using another measure. For instance, we can compute the potential profit that a market participant can realise by unwinding his position aggressively instead of unwinding his position passively. In this case, for a buy aggressive order, we compute the difference between the price $P_t$ and the best bid instead of the best ask. In fact, we tested this measure and found that the whole results in the paper hold: similar shapes for all the graphs but different figures of course. We choose to measure the potential profit at the best ask and not at the best bid in order to respect as much as possible the behaviour of market makers who are not supposed to be 100% aggressive. Their average aggressive passive ratio is around 50%, see Megarbane *et al.* (2017). Measuring the potential profit at the best ask reproduces as its best this behaviour proper to market makers.

We will also need to measure the price impact of an individual aggressive buy (resp. sell) order taking place at time $t$, and evaluated at time $t+h$, denoted by $PI_{t+h}$, and defined as follows:

$$PI_{t+h} = \frac{BP_{t+h} - BP_{t^-}}{S} * sign_t,$$

where $BP_{t^-}$ denotes the best ask (resp. bid) one microsecond before the buy (resp. sell) aggressive order.

Note that the price impact and the potential profit coincide for aggressive orders consuming a quantity less or equal to that present at the best limit. For orders consuming a quantity larger than that present at the best limit, to quantify the price impact at time $t+h$, one can compute the difference between the potential profit at time $t+h$ and the potential profit at one microsecond before the aggressive order. This is why in the following, we focus on the potential profit measure.

Later on, orders will be merged according to the categories defined in Section 2.2. To measure the average potential profit at time $t+h$ of a given order flow $A$, denoted by $APP^A_{t+h}$, we take the average among all aggressive orders belonging to the order flow $A$, weighted by the quantity of each aggressive order:

$$APP^A_{t+h} = \frac{\sum_{i\in A} PP^i_{t+h} Q^i}{\sum_{i\in A} Q^i}, \tag{1}$$

where $PP^i_{t+h}$ is the potential profit of the $i$th aggressive order, and $Q^i$ the quantity traded by the $i$th aggressive order.

In our analysis, when providing average results, we merge buy and sell aggressive orders. This is because the evolution of the price at the best limit following a buy aggressive order is quite symmetric to that following a sell aggressive order.

## 4. Analysis of aggressive orders with respect to consumed share

We distinguish between three different groups of aggressive orders, and we show how the price impact varies according to each group. Furthermore, we emphasise that each group of aggressive orders usually takes place in a specific LOB configuration. In addition to this, we obtain and interpret a relationship between the price profile 17 min before the aggressive order, the quantity present at the best limit just before the aggressive order, and the price impact following the aggressive order.

### 4.1. *Three different groups of aggressive orders*

We distinguish between three groups of aggressive orders:

- Partial aggressive orders: they consume less than the quantity at the best limit.
- Exact aggressive orders: they consume exactly the quantity at the best limit.
- N-limit aggressive orders: they consume more than the quantity at the best limit.

The limit order book and these groups are defined according to the visible liquidity in the order book without taking into account hidden liquidity which is well described in D'Hondt *et al.* (2004). In our sample, around 8% of aggressive orders consume hidden liquidity. Taking into account hidden liquidity is a possible choice but in this case, interpreting the group of exact and N-limit aggressive orders becomes complicated: market participants should forecast hidden liquidity in order to send exact or N-limit aggressive orders.

Furthermore, note that these groups do not take into account market fragmentation since our data covers trades on Euronext only. In fact, the same study can be conducted by taking into account all the platforms on which the CAC 40 assets are traded. In this case, the three different groups can be defined according to one consolidated order book computed based on the totality of platforms. Such a study is more complicated because it requires an access to all trades and orders data across all platforms with a very precise synchronisation across the whole platforms. This kind of study remains possible especially on recent data because of MiFid II recommendation that imposed a synchronisation across all platforms starting January 2018. Due to this recommendation, the French regulator has more synchronised trades data starting this date. The study requires orders data that can be obtained from the platform in the case of Euronext and from Thomson Reuters Tick History in the case of the rest of the platforms, even if it is less accurate than data furnished by

Table 1. Distribution of partial and exact aggressive orders across the different order flow categories.

| Order flow categories | Percentage of partial aggressive orders | Percentage of exact aggressive orders |
|---|---|---|
| Agency | 27% | 16% |
| HFT | 39% | 63% |
| Proprietary | 31% | 21% |
| RMO | 3% | 0% |

the platform directly. For an example of such a study, see Saliba (2019b).

Now that we have defined thesegroups of aggressive orders, note that exact and N-limit aggressive orders mechanically change the price since they trigger a best price change right after the trade. This is why the price impact one microsecond after these aggressive orders is obviously significant. We aim at investigating whether these mechanical impacts are temporary or reflect a certain information content that persists over time.

### 4.2. *Some preliminary statistics*

In our database, partial and exact aggressive orders constitute the majority of aggressive orders (96%). Furthermore, HFTs send more exact aggressive orders than partial ones: 63% of the exact aggressive orders are sent by HFTs, while only 39% of partial ones are sent by them (see Table 1).

On average, one partial aggressive orders consume a volume (11 k €) almost equal to that consumed by exact aggressive orders (13 k €). N-limit aggressive orders consume an amount clearly more significant than the other aggressive orders (43 k €). It is important to point out that aggressive orders, and especially exact (resp. N-limit) aggressive orders occur upon particular conditions: when the quantity at the best limit is significantly less than the average quantity at this limit over all events: 13 k € (resp. 16 k €) just before the exact (resp. N-limit) aggressive orders, versus 57 k € on average (see Table 2).

This is not really surprising, and is related to the well-known information content of the order book imbalance: when the quantity on one side of the book is significantly larger than that on the other side, the next aggressive order will likely hit the smallest side, see for example Stoikov (2018). We now investigate more precisely the relationship between the imbalance and the aggressive order group.

### 4.3. *Relationship between imbalance and aggressive order group*

We show in this section that market participants submit exact aggressive orders when the LOB is significantly imbalanced. The imbalance at time $t$, just before the aggressive buy (resp. sell) order takes place, is computed as follows:

$$Imb_t = \frac{Q_t^1 - Q_t^2}{Q_t^1 + Q_t^2},$$

where $Q_t^1$ denotes the quantity present at the best bid (resp. ask) at time $t$, and $Q_t^2$ denotes the quantity present at the best ask (resp. bid) at time $t$ when it is a buy (resp. sell) aggressive order.

The value of the imbalance one microsecond before the exact aggressive trades (on average equal to 27%) is significantly higher than that (on average equal to 3%) before the partial aggressive trades (see Figure 2).

### 4.4. *Price impact according to the groups of aggressive orders*

As expected, one microsecond after the aggressive order, because of the mechanical impact, the price impact due to N-limit aggressive orders is higher than that of exact ones, which is higher than that of partial ones (see the price profiles in Figure 3, from which price impacts are obviously deduced).

One relevant question is whether this mechanical impact is permanent or not. Figure 3 shows that the price impact of exact aggressive orders is permanent: it is higher than that of partial ones, over all time horizons, almost equal to two-thirds of the bid-ask spread. On the contrary, N-limit aggressive orders have a temporary component in their price impact: market participants tend to refill the LOB by submitting new orders in place of the consumed ones. Indeed, starting one second after the aggressive order, the price impact begins to attenuate. On a 17 min time horizon, the remaining mechanical impact of N-limit aggressive orders is quite equal to that of exact aggressive orders (recall that Figure 3 displays the price profiles, and that the price impact of N-limit aggressive orders is deduced as the difference between the profile at time $t$ and Point A).

In order to study the robustness of the results of Figure 3, we repeat the same experience, but this time by computing the average measure month by month instead of considering the whole period. Figure 4 gives the corresponding results.

Figure 4 shows that the results relative to partial and exact aggressive orders are robust, while that of N-limit aggressive orders is less robust. In September, the price impact of the

Table 2. General statistics on the different groups of aggressive orders.

| | Average traded amount per aggressive order | Median traded amount per aggressive order | Share of aggressive orders number | Share of traded amount | Amount at the best limit just before the aggressive order |
|---|---|---|---|---|---|
| Partial aggressive orders | 11 k € | 6 k € | 49.5% | 38% | 41 k € |
| Exact aggressive orders | 13 k € | 8 k € | 46.5% | 48% | 13 k € |
| N-limit aggressive orders | 43 k € | 22 k € | 4% | 14% | 16 k € |

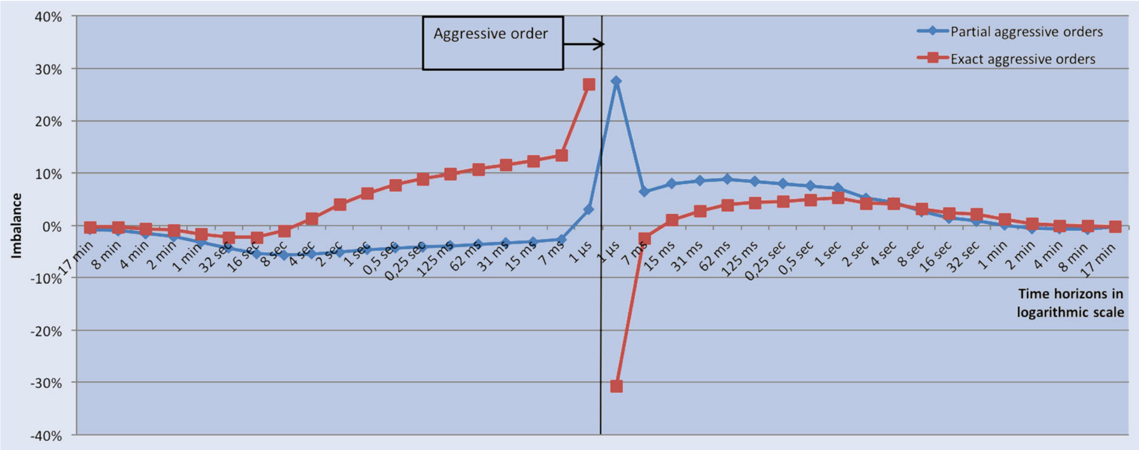


Figure 2. Variation of the imbalance of the LOB before and after the arrival of partial and exact aggressive orders.

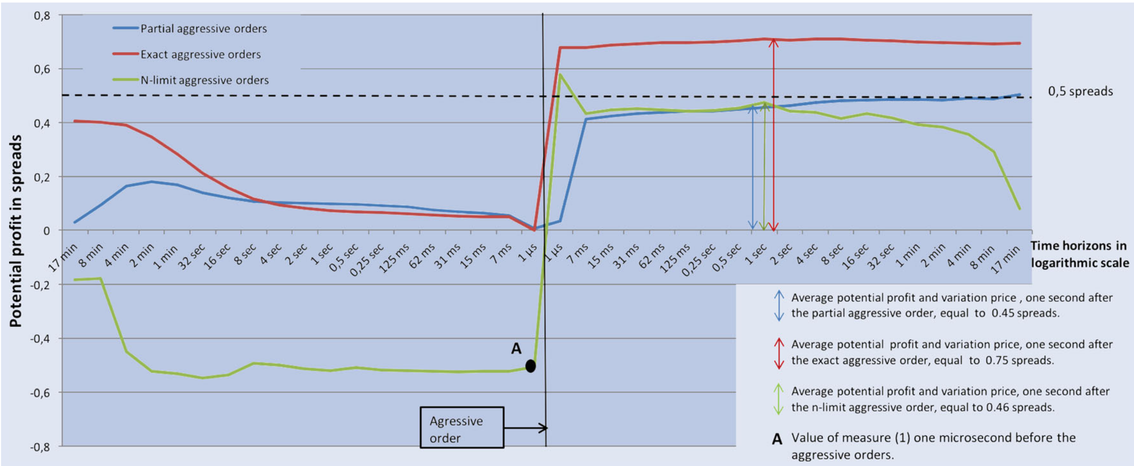


Figure 3. The price profile measure (1) according to the aggressive order groups.

N-limit aggressive orders starts to decrease starting one second after the aggressive orders, while it is permanent during October and November. This shows that the result related to N-limit aggressive orders found in Figure 3 are due to noise.

### *4.5. Focus on partial and exact aggressive orders*

We deepen our analysis by investigating the price profile of partial (consuming less than 100% of the best limit) and exact (consuming 100% of the best limit) aggressive orders according to the consumed share at the best limit in percentage. In general, the price impact is studied according to the traded volume. In this work, we choose on purpose to study it according to the consumed share in order to show that the price impact does not only depend on the traded volume but also on the quantity present at the best limit. In particular, we want to understand the relationship between historical prices evolution, consumed share and price impact.

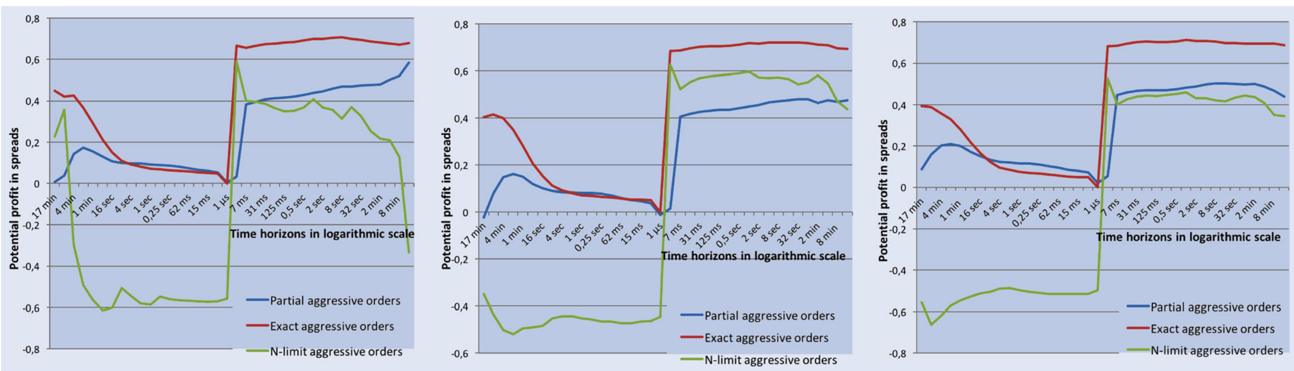


Figure 4. The price profile measure (1) according to the aggressive order groups computed respectively on September, October and November.

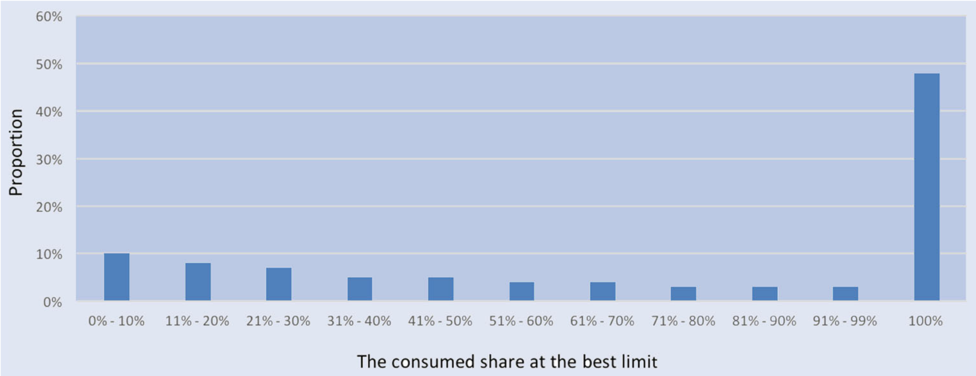


Figure 5. Distribution of partial and exact aggressive orders according to the consumed share at the best limit.

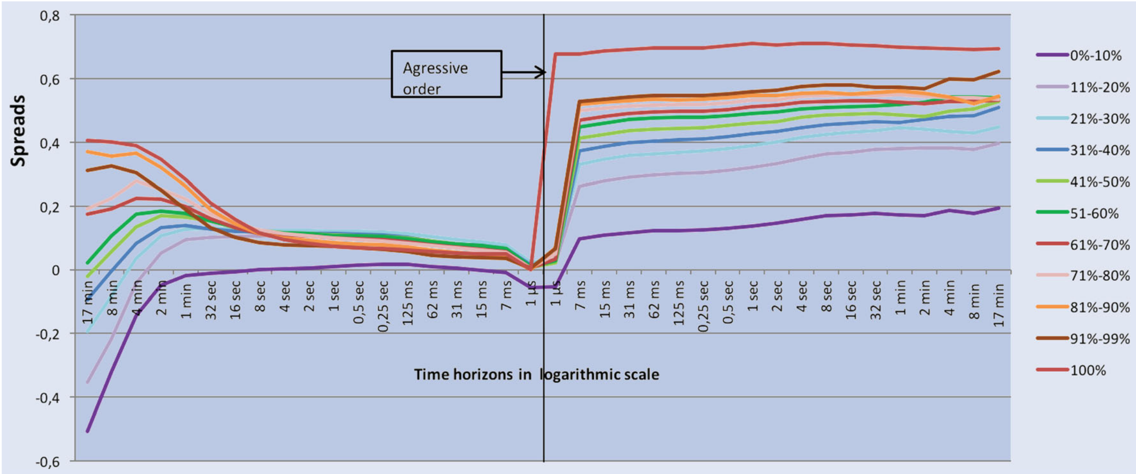


Figure 6. The price impact according to the consumed share at the best limit.

In Figure 5, we show the proportion of aggressive orders according to the consumed share at the best limit (N-limit aggressive orders excluded).

In Figure 6, we plot the price impacts according to the consumed share.

Figure 6 shows that the magnitude of the price impact over all time horizons after the aggressive order is increasing with respect to the consumed share. As an example, the price impact following an aggressive order consuming 10% of the total quantity present at the best limit is significantly lower than that due to an aggressive order consuming 90% of the quantity present at the best limit. This could be interpreted by the fact that the imbalance created following an aggressive order consuming 90% of the quantity at the best limit is higher than the one following an aggressive order consuming only 10%. As already seen, a large imbalance is likely to trigger other aggressive orders (or cancellations of limit orders).

We now investigate whether the consumed share rather depends on the quantity present at the best limit or the traded amount. Figure 7 shows that as expected, the consumed part varies with the traded amount, but also depends significantly on the quantity present at the best limit. This means that the

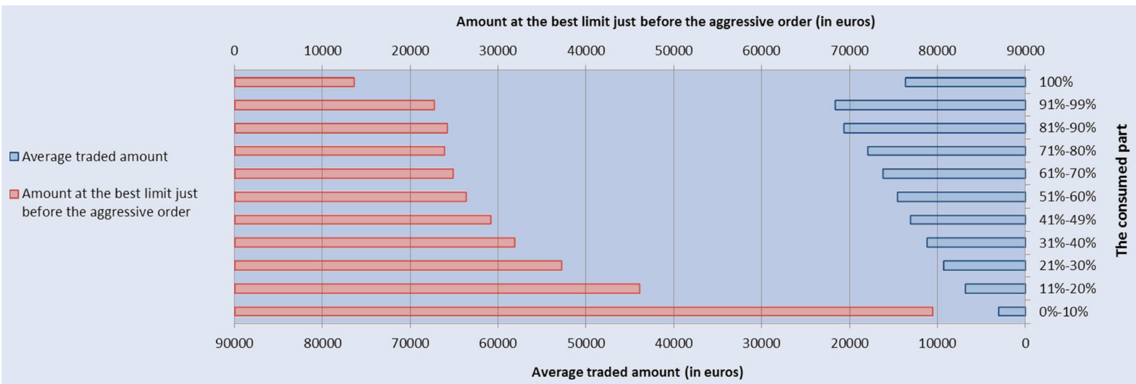


Figure 7. Evolution of the amount at the best limit just before the aggressive orders and the average traded amount according to the consumed part.

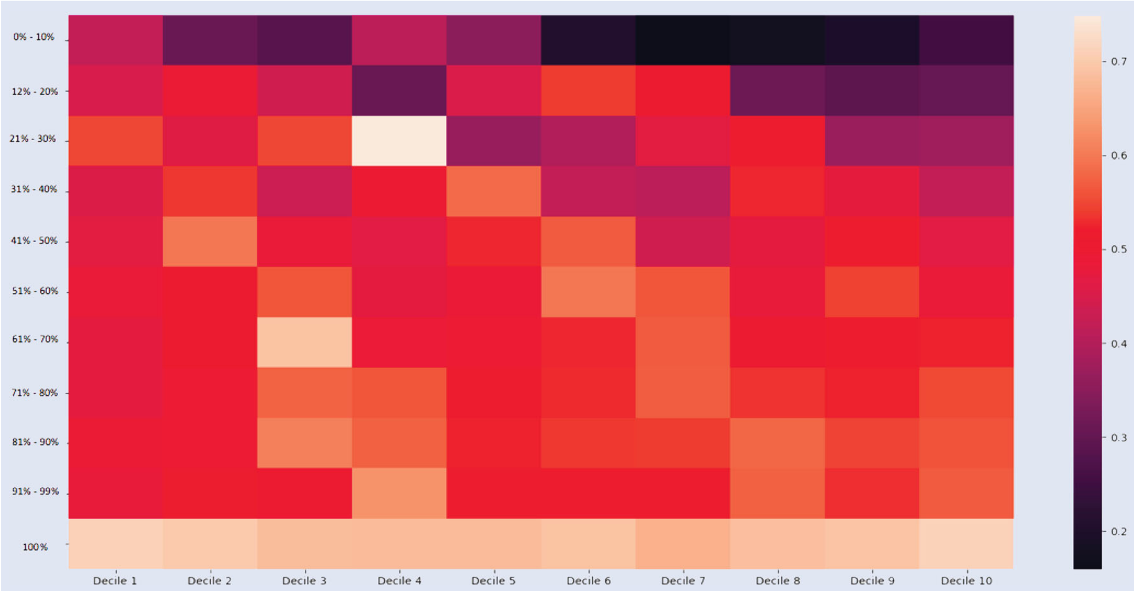


Figure 8. The heatmap of the price impact vs the consumed fraction (the $y$-axis) and the consumed volume (the $x$-axis).

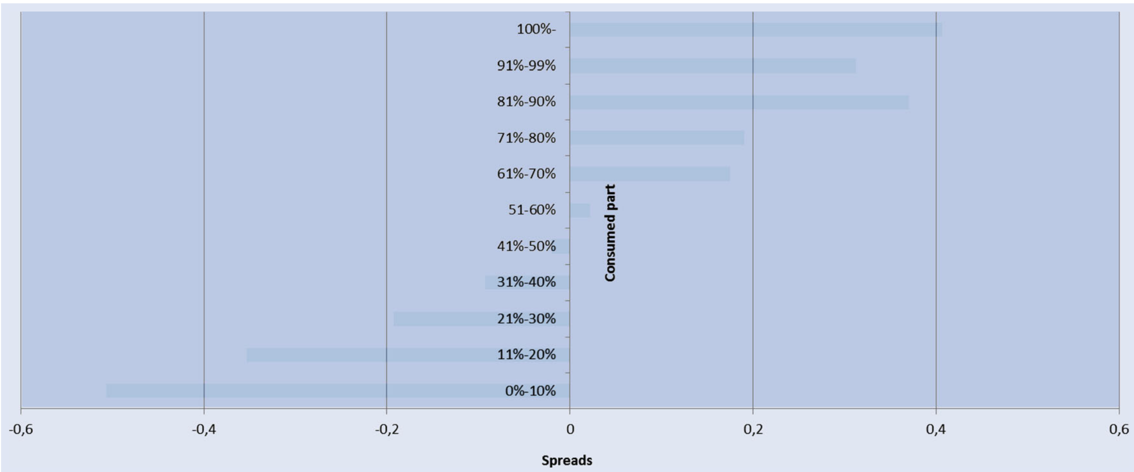


Figure 9. Evolution of the consumed part at the best limit according to the measure (1) evaluated 17 min before the aggressive order.

price impact does not depend only on the traded volume, but on the volume present at the best limit too, which is showed more clearly in Figure 8.

Figure 8 shows that the consumed fraction explains better the price impact variation than the traded volume: when observing the colour variation in the graph, it is clear than the colour gets darker on the top and lighter on the bottom, which means that the price impact increases with the consumed fraction. This colour variation is less clear when observing the colour variation vertically, according to the consumed fraction represented by the $x$-axis.

Another important feature appearing in Figure 6 is that the consumed share is increasing with respect to the price profile measure (1) evaluated 17 min before the aggressive order. To clarify this phenomenon, we plot in Figure 9 the measure (1) evaluated 17 min before the aggressive order with respect to the consumed part at the best limit.

As previously shown, the consumed share depends on the quantity present at the best limit just before the aggressive order, which depends on the historical evolution of the price. For instance, if the price has been decreasing, the quantity present at the best ask tends to be small, since new limits are revealed. Following this logic, the quantity present at the best limit is a decreasing function with respect to the price profile 17 min before the aggressive order, which explains the relationship appearing in Figure 9.

## 5. Potential profits according to the different order flow categories

We study in this section the potential profits according to each order flow category after partial and exact aggressive orders. We then focus on partial aggressive orders, in order to identify the potential profit disparities between market participants within the same order flow category. Finally, for a given market participant having activities belonging to different order flow categories, we check the differences in potential profit according to these different categories.

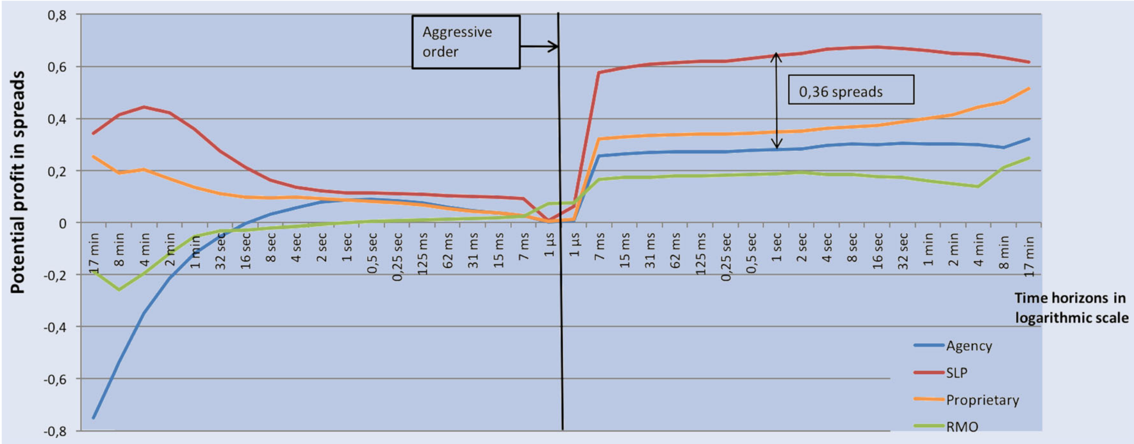


Figure 10. The potential profit evolution following partial aggressive orders according to the different order flow categories over different time horizons.

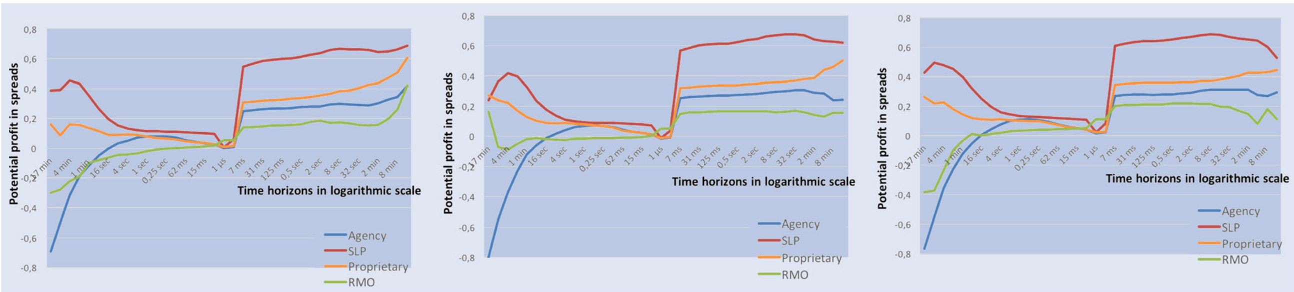


Figure 11. The potential profit evolution following partial aggressive orders according to the different order flow categories over different time horizons computed respectively on September, October and November.

### *5.1. Potential profits after partial aggressive orders*

The HFT flow stands out with the (significantly) highest potential profit in the case of partial aggressive orders, over all time horizons. One second after partial aggressive orders, HFTs have a potential profit 0.36 spreads higher than agency participants, and 0.29 spreads higher than proprietary participants (see Figure 10). RMO members are the least profitable. In addition to this, we show in Section 6.1 that the aggressive orders of HFTs are the less autocorrelated, which allows us to deduce that the high potential profit of HFTs is due to an informational advantage and not to autocorrelated orders.

In order to study the robustness of the results of Figure 10, we repeat the same experience, but this time by computing the average measure month by month instead of considering the whole period. Figure 11 gives the corresponding results.

Figure 11 shows that the results of Figure 10 are robust, this is because the results obtained for the three months are similar.

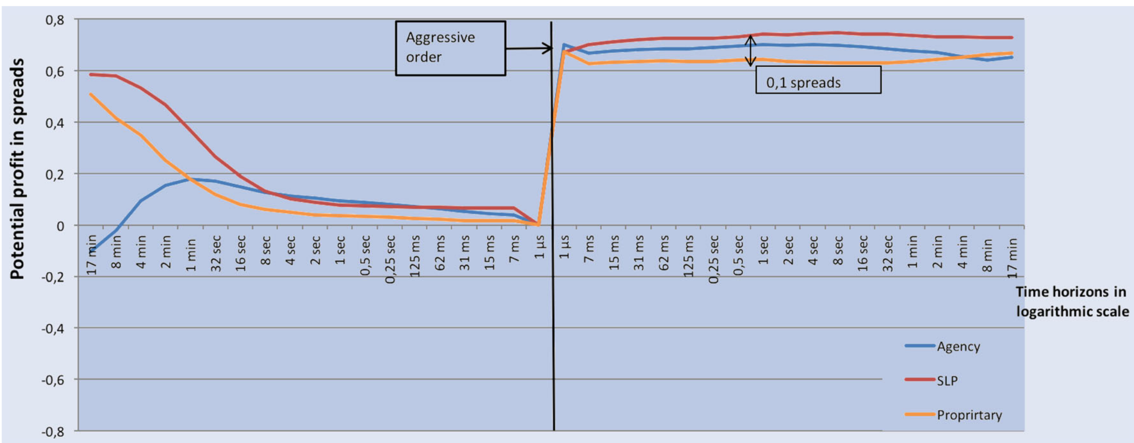


Figure 12. The potential profit evolution following exact aggressive orders according to the different order flow categories over different time horizons.

Table 3. Number of member codes issuing enough partial aggressive orders according to each order flow category.

| Order flow category | Number of member codes | Number of member codes issuing enough partial aggressive orders |
|---|---|---|
| Agency | 74 | 33 |
| SLP | 17 | 11 |
| Proprietary | 72 | 24 |
| RMO | 23 | 6 |
| RLP | 4 | 2 |

### 5.2. *Potential profits after exact aggressive orders*

Although HFTs still obtain a better potential profit than other market participants in the case of exact aggressive orders (see Figure 12), the difference between the categories is not much significant: the potential profit of HFTs is only 0.04 spreads higher than that of agency members and 0.1 spreads higher than that of proprietary members (see Figure 12).

In order to study the robustness of the results of Figure 12, we repeat the same experience, but this time by computing the average measure month by month instead of considering the whole period. Figure 13 gives the corresponding results.

Figure 13 shows that the results of Figure 12 are robust, this is because the results obtained for the three months are similar.

Furthermore, Figures 10 and 12 show that on average, HFTs and proprietary members are mean reverting: by this, we mean that they buy when the price decreases and sell when the price increases. In contrast, agency members seem on average trend following: they buy when the price increases and sell when the price decreases. These results are in line with the finding in Biais *et al.* (2016).

### 5.3. *Does the potential profit vary among members within the same order flow category?*

We now investigate the potential profit disparities between different members belonging to the same order flow category for partial aggressive orders.

Table 3 shows the total number of member codes according to each order flow category and the number of member codes issuing enough† partial aggressive orders.

We compute the proportion of member codes with a potential profit higher than the third quartile (based on the potential profits of member codes having enough partial aggressive orders) for each order flow category over different time horizons following the partial aggressive order. These proportions are plotted in Figure 14.

In order to understand the results in Figure 14, we take the following example: the value relative to HFT activity 7 ms after the aggressive order is equal to 90%. This means that 90% of the HFTs member codes have a potential profit higher than the third quartile. Now we can interpret the rest of the results. We find that over a short time horizon (until 2 min after the aggressive order), HFTs belong to the 25% market participants realising the highest short-term potential profits. Over a longer time horizon, from two minutes after the aggressive order, the proportion of HFTs with potential profit higher than the third quartile starts to decrease to the benefit of proprietary traders (see Figure 14). This could be due to the fact that HFTs do not target ‘ long-term’ strategies, high-frequency trading being an activity where participants typically hold positions for a very short times.

### 5.4. *Disparities in potential profits for a same member code according to the different order flow categories*

We now dissociate the flows of a same member code according to the order flow categories. It allows us to identify the different potential profits generated by a member. We can notably distinguish between high-frequency trading strategies targeting short-term potential profits and longer-term strategies.

For members carrying out simultaneously SLP and another activity, the potential profit of the SLP flow is always higher than that of the other flows. In the majority of cases, the potential profit of the proprietary flow is higher than or equal to that of the agency flow. We illustrate these findings through two examples in Figures 15 and 16.

In Figure 15, we plot the different potential profits of Member code A (having SLP and proprietary activities at the same time) according to the order flow categories.

Until 32 s after the aggressive order, the SLP flow has a potential profit higher than the third quartile, while the potential profit of the proprietary flow is equal or lower than the first quartile. The proprietary activity of Member code A seems to target a longer-term strategy: 2 min after the aggressive order, its potential profit becomes higher, outperforming the one of SLP. On a 17 min horizon, it is equal to 2.3 spreads, 3 times higher than the SLP flow potential profit.

In Figure 16, we plot the different potential profits of Member code B, according to the order flow categories.

Member code B has different potential profit levels, depending on the order flow category considered. The potential profit of the SLP flow is the highest at any time scale: one second after the aggressive order, the potential profit of the SLP flow (0.43 spreads) is higher than that of the proprietary one (0.37 spreads) and significantly higher than the agency flow (0.27 spreads).

## 6. From aggregated aggressive orders to strategies

In this section, we show that the analysis of aggressive orders is useful to understand other features than price impact and potential profit. For instance, we study the autocorrelation of the different order flow categories. In addition to this, we propose a new classification of member codes (whose flows are segmented according to the order flow category) based on the investigation of aggressive orders. We also show that we can access to a more granular classification by segmenting member code flows according to the different connectivity channels they use. Finally, by observing the evolution of the

† We consider that the number of aggressive orders is enough when there is at least one aggressive order per day and per asset.

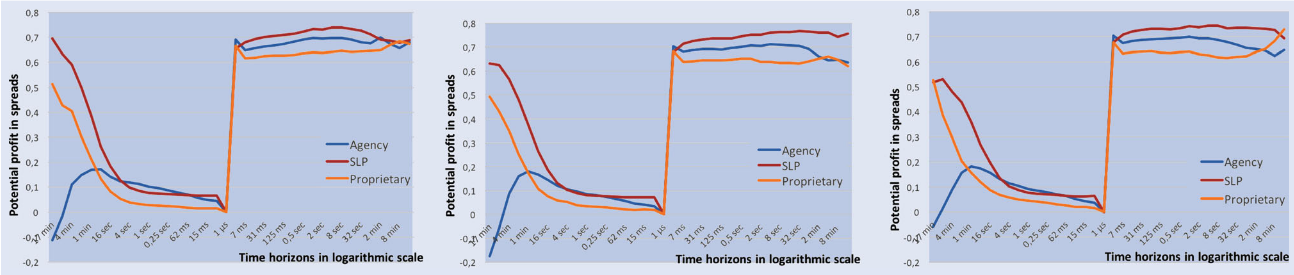


Figure 13. The potential profit evolution following exact aggressive orders according to the different order flow categories over different time horizons computed respectively on September, October and November.

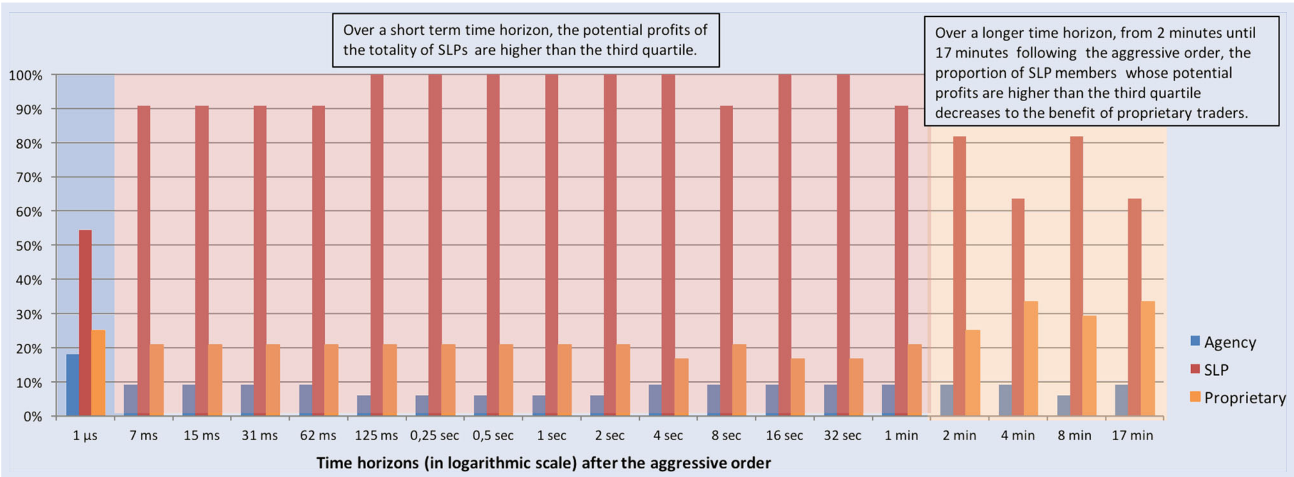


Figure 14. Evolution of member codes proportion with a potential profit higher than the third quartile for each order flow category over different time horizons following the partial aggressive order.

price before the aggressive order takes place, we deduce the different strategies of member codes, such as mean reverting or trend following.

### 6.1. *Autocorrelation of aggressive orders according to the different order flow categories*

We study the autocorrelation properties of each order flow category, taken separately. Positive autocorrelation means that once an aggressive order has been observed, the probability to observe another one in the same direction is larger than one half. Typically if an agency broker is splitting a large client's metaorder in slices, its flow exhibits a significant positive autocorrelation, see Bouchaud *et al.* (2009) and Toth *et al.* (2015).

Figure 17 shows that the autocorrelation of the first two successive aggressive orders is almost equal among all order flow categories (equal respectively to 29%, 30% and 34% for agency, proprietary and SLP order flow categories). The autocorrelation of the aggressive orders of the agency order flow category is the highest among all order flow categories (with the exception of the first two aggressive orders). The decrease

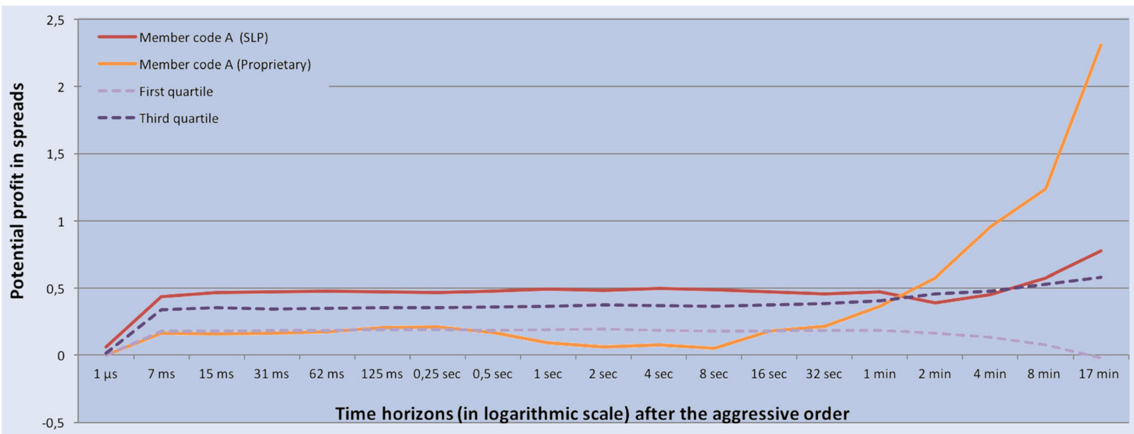


Figure 15. Disparities in potential profits of Member code A according to the different order flow categories issued by this member code. The quartiles are the same as those computed in Section 5.3.

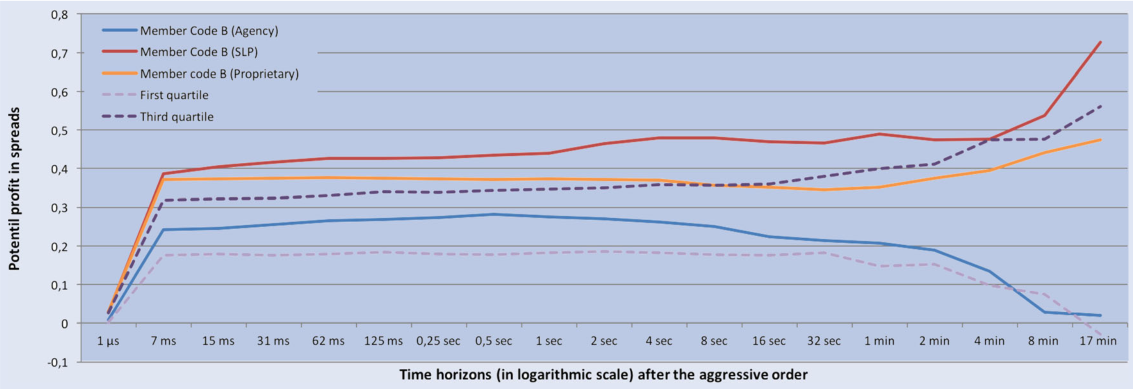


Figure 16. Disparities in potential profits of Member code B according to the different order flow categories issued by this member code.

in autocorrelation is the most significant for the SLP order flow category. The aggressive orders autocorrelation in this flow almost vanishes from the 4th aggressive order (it is equal to 5% in this case).

### *6.2. A classification tool*

It is usual to consider passive orders to classify members as HFTs or non-HFTs. One of AMF classifications is based on this type of orders.

This AMF classification differentiates between three classes of market participants: HFTs, mixed HFTs and non-HFTs. It is based on the lifetime of cancelled orders and determined using two sets of conditions:

- **Condition 1 is based on a comparison with other participants:** the participant must have cancelled at least 100,000 orders during the year, and the average lifetime of his cancelled orders should be less than the average lifetime of all cancelled orders in the book.
- **Condition 2 is based on a set threshold:** the participant must have cancelled at least 500,000 orders with a lifetime of less than 0.1 second (i.e. the participant quickly updates the orders in the limit order book) and the top percentile of the lifetime of its cancelled orders must be less than 500 ms (i.e. the participant regularly uses fast access to the market).

A member code is a high-frequency trader if it is not an investment bank and it meets one of these conditions. An investment bank meeting one of these conditions is described as mixed HFT. Note that some members satisfy Condition 2 without satisfying Condition 1.

However, it seems also possible to classify participants by relying on aggressive order potential profits. Those realising the higher short-term potential profits (one second after the aggressive order) can be considered as HFTs, and those realising the lowest as non-HFTs. It turns out that relying on both approaches allows us to obtain a more complete classification of market participants. Three different classes can be distinguished for member codes whose flows are segmented according to the order flow category:

- Pure HFTs are characterised by a high short-term potential profit (higher than the third quartile of potential profits computed among all market participants) and a low lifetime of cancelled orders (lower than the third quartile of lifetimes of cancelled orders computed among all market participants).
- Pure non-HFTs are characterised by a high lifetime of cancelled orders and a small short-term potential profit.
- Intermediary agents are characterised by a small short-term potential profit and a low lifetime of cancelled orders.

We point out that, as expected, no member code has high short-term potential profits and high lifetime of cancelled orders (see Figure 18). Moreover, note that all SLPs belong the pure HFT category.

### *6.3. A more granular classification using the different connectivity channels*

Market members connect to Euronext via connectivity channels (called ‘SLE’, a French acronym for ‘Serveur Local d’Emmission’, since this technology was initially developed by Canadian and French companies)† to convey their orders. Some member codes use their SLEs to separate their aggressive flow from their passive one, and some others to separate their different order flow categories. Dissociating the flows issued by a same member code, and belonging to the same order flow category according to SLEs can in some cases bring up new information concerning the different activities followed by this member code. The main advantage of using different SLEs is that it is easy to setup specific sets of risk limits for each of them (like the number of orders per day or per hour, the maximum traded value, etc). For instance, we dissociate the flow of Member code B who is an agency

† On the three studied months and over the CAC 40 stocks, there are in total 355 SLEs. 94% of SLEs are used by only one member code, 6% are used by two member codes (belonging to the same institution). 73% of SLEs are deployed for only one order flow category (Agency, Proprietary, HFT, RMO), 26% are deployed for two different order flow categories and 1% for three different order flow categories. The number of SLEs belonging to the same institution varies between 1 and 33. There are 46 member codes using more than one SLE over 117 member codes in total.

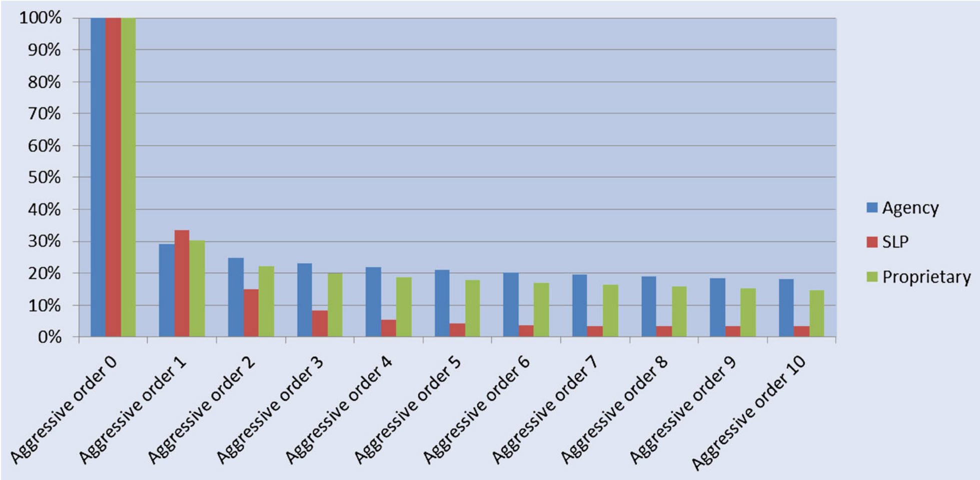


Figure 17. Autocorrelogram of aggressive orders according to each order flow category.

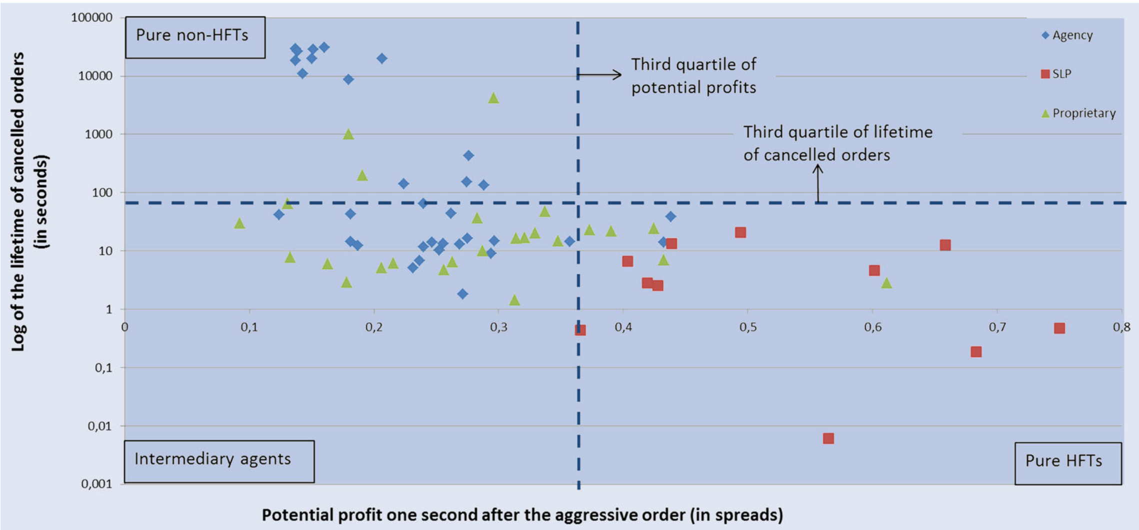


Figure 18. Using both methods of classification: one relying on cancellations and one relying on potential profits (the flows are grouped by member code and order flow category).

broker serving as an intermediary for a HFT (among other clients) according to the different SLEs. We plot in Figure 19 the potential profit of each of these flows.

The difference in potential profits between the two flows, illustrated in Figure 19 could be interpreted for example by a segmentation of these different flows allowing us to identify according to the clients' typology: one SLE is dedicated for a HFT client while another is dedicated for other type of clients.

### 6.4. *Different strategies*

By observing what happens before the aggressive order, we can distinguish between three different strategies:

- Mean reverting strategy, going against the price variations (see Figure 20).
- Trend following strategy, following the price variations (see Figure 21).
- Another strategy consisting in benefiting from the insertion of new orders that reduce the spread (see Figure 22).

In Figures 20 and 21, we pick respectively some mean reverting and trend following SLPs and we plot their price profiles (using Measure (1)) around their aggressive orders.

In Figure 22, we plot the price profiles of three HFTs following a particular strategy consisting in seising certain opportunities faster than other market participants (Institution 1, Institution 2 and Institution 3) and another HFT (Instituton 10) who does not follow this same particular strategy.

The arbitrage configuration in Figure 22 clearly stands out: the three institutions benefit from the insertion of new orders inserted in the LOB 7 ms before the aggressive order reducing the spread by 0.6 spreads on average (in the case of Institution 3) and by 0.2 spreads on average (in the case of Institutions 1 and 2). These situations are likely to occur when the spread is large (equal to several ticks). In order to check that this arbitrage configuration does not take place only at the beginning or the end of the day, we performed the same computations by excluding all the aggressive trades taking place before 11h and after 16h. The obtained results are quite unchanged. We deduce then that Institutions 1, 2 and 3 do not apply this strategy particularly at the beginning or at the end of the day. This

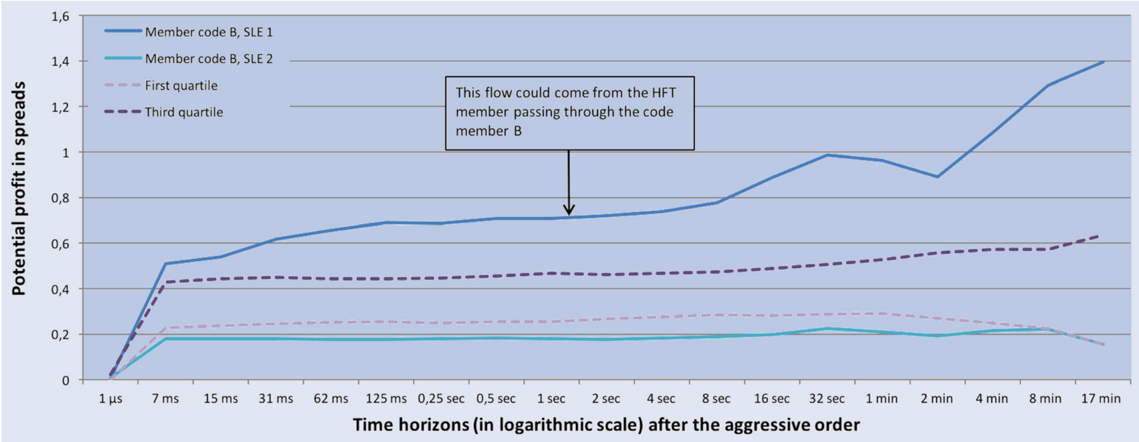


Figure 19. Disparities in potential profits of the same agency broker according to its different SLEs.

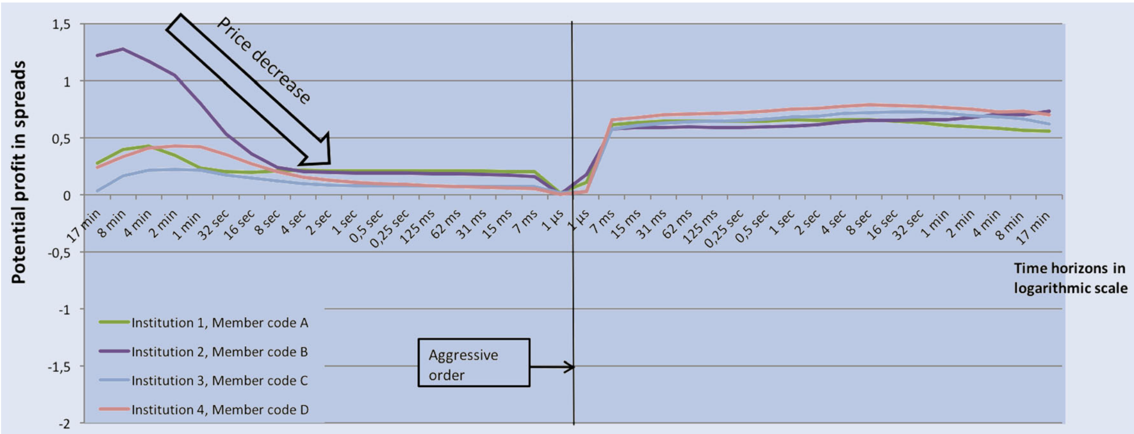


Figure 20. Mean reverting HFTs.

arbitrage configuration might not be a real arbitrage configuration, but a consequence of a scenario where these institutions perform instantaneous wash trades: they insert the limit order inside the spread and then consume their own limit orders. To make sure that we are not in this kind of situation, we performed the same computations by excluding all the wash trades of all institutions. The obtained results are unchanged. We deduce that this is a real arbitrage configuration and not a wash trades one.

Some institutions carry out distinct strategies simultaneously. For instance, Institution 1 follows mean reversion and trend following strategies at the same time using different member codes. By using the same member code used for the mean reversion strategy, Institution 1 succeeds to profit from local opportunities by benefiting from the insertion of new orders that reduce the spread (it is the particular strategy showed in Figure 22).

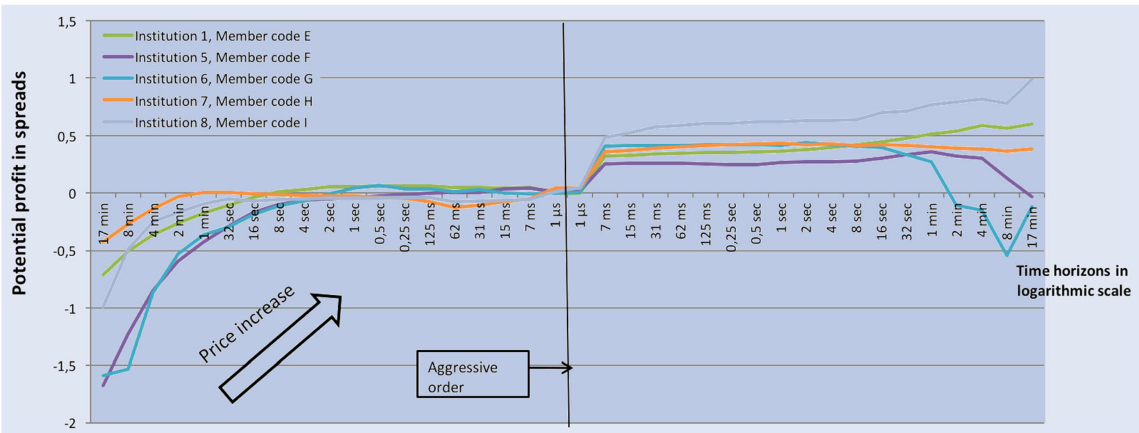


Figure 21. Trend following HFTs.

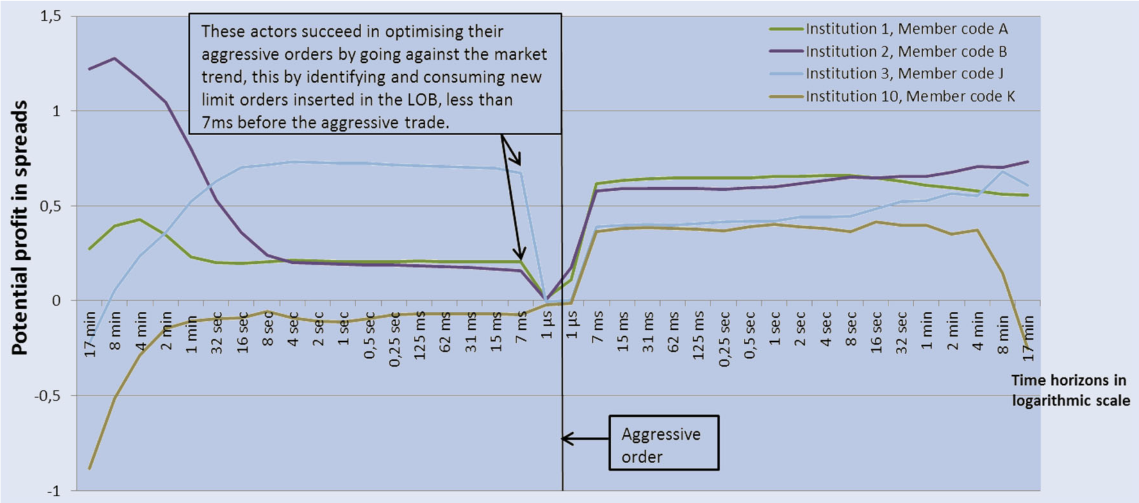


Figure 22. Some HFTs following a particular strategy consisting in benefiting from the insertion of new orders that reduce the spread.

## 7. Conclusion

We show that the proportion of executed volume at the best limit does have a significant and durable impact on prices: the price impact depends on both traded volume and volume present at the best limit. In terms of strategy, when observing the price evolution before the aggressive order, it is clear that HFTs are typically mean reverting, while agency members are trend following. In addition to this, we have seen that the aggressive orders of HFTs are more informed than those of other market participants. Finally, we have validated some expected stylised facts: The aggressive orders of agency members are the most autocorrelated, while those of HFTs are the least autocorrelated. This shows that the high potential profit of HFTs is not due to the price impact that they generate but to a real informational advantage. This informational advantage is mainly due to the fact that HFTs use more sophisticated infrastructure and automation technologies than other participants, which allows them to predict the price evolution before the rest of market participants. These findings can explain why some exchanges plan to reduce the aggressive behaviour of HFTs by imposing 'speed bumps' which aim is to impose a speed limit on their aggressive trades. These speed bumps that have mostly been introduced in American exchanges recently start to become more popular. This is the case for example of the Deutsche Börse Eurex platform that will test a 6-month pilot project from June 3, 2019 slowing down the HFTs aggressive orders by one millisecond for trading on German and French options.


## Acknowledgments

We are grateful to Alexandra Givry, Philippe Guillot, Charles-Albert Lehalle, Julien Leprun, Nicolas Megarbane and Mathieu Rosenbaum for their valuable comments.


## Disclosure statement

No potential conflict of interest was reported by the author(s).

## References


Ait-Sahalia, Y. and Sağlam, M., High frequency market making: Optimal quoting, 2017. Available at SSRN 2331613.

Almgren, R., Thum, C., Hauptmann, E. and Li, H., Direct estimation of equity market impact. *Risk*, 2005, **18**(7), 58–62.

Biais, B., Declerck, F. and Moinas, S., Who supplies liquidity, how and when?, 2016.

Bouchaud, J.-P., Farmer, D. and Lillo, F., How markets slowly digest changes in supply and demand. In *Handbook of Financial Markets: Dynamics and Evolution*, pp. 57–160, 2009 (Elsevier: Amsterdam).

Brogaard, J., Hendershott, T. and Riordan, R., High-frequency trading and price discovery. *Rev. Financ. Stud.*, 2014, **27**(8), 2267–2306.

Cardaliaguet, P. and Lehalle, C.-A., Mean field game of controls and an application to trade crowding. *Math. Financ. Econ.*, 2017, **12**, 335–363.

Chordia, T., Roll, R. and Subrahmanyam, A., Market liquidity and trading activity. *J. Finance*, 2001, **56**(2), 501–530.

D'Hondt, C., De Winne, R. and Francois-Heude, A., Hidden orders on euronext: Nothing is quite as it seems, 2004. Available at SSRN 379362.

Euronext. Revision of the supplemental liquidity provider programme, 2016.

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

Jones, C.M., What do we know about high-frequency trading? Columbia Business School Research Paper No. 13-11, 2013.

Jovanovic, B. and Menkveld, A.J., Middlemen in limit order markets. Technical report, SSRN, 2016.

Lehalle, C.-A. and Mounjid, O., Limit order strategic placement with adverse selection risk and the role of latency. *Market Microstruct. Liquid.*, 2017, **3**(1), 1750009.

Megarbane, N., Saliba, P., Lehalle, C.-A. and Rosenbaum, M., The behavior of high frequency traders under different market stress scenarios. *Market Microstruct. Liquid.*, 2017, **3**(3–4), 1850005.

Menkveld, A.J., High frequency trading and the new market makers. *J. Financ. Markets*, 2013, **16**(4), 712–740.

Moro, E., Vicente, J., Moyano, L.G., Gerig, A., Farmer, D., Vaglica, G., Lillo, F. and Mantegna, R.N., Market impact and trading profile of hidden orders in stock markets. *Phys. Rev. E*, 2009, **80**(6), 066102.

Saliba, P., Analysis of the aggressive behaviour of market participants: Do HFTs trade opportunely? Technical report, AMF, 2019a.

Saliba, P., High-frequency trading: Statistical analysis, modelling and regulation. PhD Thesis, 2019b.

Stoikov, S., The micro-price: A high frequency estimator of future prices. *Quant. Finance*, 2018, **18**(12), 1959–1966.
Tóth, B., Lemperiere, Y., Deremble, C., De Lataillade, J., Kockelkoren, J. and Bouchaud, J.-P., Anomalous price impact and the critical nature of liquidity in financial markets. *Phys. Rev. X*, 2011, **1**(2), 021006.
Toth, B., Palit, I., Lillo, F. and Farmer, J.D., Why is equity order flow so persistent? *J. Econ. Dyn. Control*, 2015, **51**, 218–239.

## Appendix. Generalisation to all HFTs

We generalise the study of partial aggressive orders to all HFT members (and not to the members of the SLP programme only) that we classify according to the classification described in Section 6.2 based on the lifetime of cancelled orders. We compare their potential profit to those of other market participant classes (mixed members and non-HFTs).

The Table A1 shows the total number of member codes and the number of member codes issuing enough partial aggressive orders (using the same criterion as in Section 5.3) relative to each market participant class. The results here are quite similar to those obtained for SLPs in the previous Section 5.3, (see Figure A1).

Out of the 12 remaining HFTs, 8 are SLPs.

Figure A1 shows that over a short time horizon (from 7 milliseconds until two minutes approximately after the aggressive trade), the majority of HFTs (between 77% and 92% of them) display a potential profit higher than the third quartile, versus 18% on average (the average value is computed starting 7 milliseconds until 1 minute after the aggressive order) for mixed members and only 3.5% for non-HFTs. Beyond one minute, the presence of HFTs over the third quartile decreases to the benefit of other participants, in particular the mixed member codes. From 31 ms to 4 s after the aggressive order, the proportion of HFTs having a potential profit higher than the third quartile is quite constant, equal to 92%. During this time interval, one HFT member code (whose aggressive flows constitute 7% of the total aggressive flows of this institution) only does not realise potential profits higher than the third quartile. We note that the other member codes of this institution are more profitable than 75% of the market participants over the studied time horizon.

Table A1. Number of member codes issuing enough partial aggressive orders according to each market participant class.

| Market participant class | Number of member codes | Number of member codes issuing enough partial aggressive orders |
|---|---|---|
| HFT | 20 | 12 |
| Mixed | 13 | 13 |
| non-HFT | 85 | 30 |

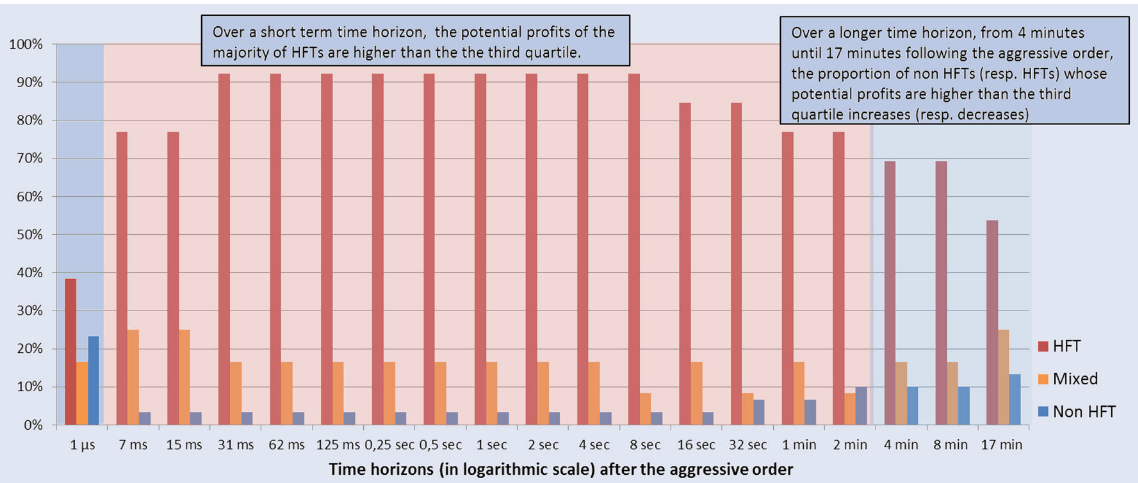


Figure A1. Evolution of member codes proportion with a potential profit higher than the third quartile for each market participant class over different time horizons following the partial aggressive order.