Optimal high frequency strategy in an omniscient order book

Marouane Anane, Frédéric Abergel

Optimal High Frequency Strategy in Omniscient Order Book

Marouane ANANE [†*] Frederic ABERGEL [†‡]

February 5, 2014

Abstract

The aim of this study is to quantify the low latency advantage of High Frequency Trading (HFT) and to compute, empirically, an optimal holding period of a HF trader. Critics claim that low latency leads to information asymmetry victimizing retail investors. However, objective studies measuring the gain due to this asymmetry are rare. In order to perform the study, new methods are introduced in this paper, in particular, the optimal strategy problem is formulated and ideas are given to compute it in a reasonable amount of time. A new measure, the weighted mean holding period, is introduced and an algorithm to compute it is suggested. Using the previous concepts, a large empirical study based on optimal omniscient strategy is presented and evidence of the low latency advantage limitation is provided. In particular, it is shown that the bid ask spread and the transaction costs lead to a trading frequency much lower than the information renewal frequency.

Keywords : High Frequency Trading, Omniscient Order Book, Optimal Strategy, Holding Period, Linear Programming, Sparse Matrices.

Introduction

Since the last financial crisis, proprietary trading, especially High Frequency Trading, has been widely criticized and assumed to be one of the market instability main causes. In 2010, President Obama's adviser has argued [3] that such speculative activity played a key role in the financial crisis of 2007-2010. Many regulation ideas have been suggested. Tobin Tax [11] is a well-known example.

The rationale behind penalizing HFT agents is to protect investors from such professional speculators. HFT firms are widely assumed to be armed with sophisticated mathematical algorithms and a strong software framework allowing them to make big profits by rapidly making the best decisions. Due to the short holding period, HFT seems to be a risk-free activity providing huge profits, victimizing less sophisticated investors. HFT is also assumed to cause flash crashes, artificial volatility, and to increase market adverse selection by hitting the order book systematically at each arbitrage opportunity.

[*]marouane.anane@ecp.fr

[†]frederic.abergel@ecp.fr

[‡]Chaire de Finance Quantitative, Laboratoire MAS, Ecole Centrale Paris, Grande voie des vignes, 92290 Chatenay Malabry, France

Despite all these assumptions, empirical papers published by different authors studying the US market claim modest profit's upper bound. Kearns, et al. [7] demonstrated that HFT profits are modest compared to the traded volume. In particular, their study found an upper bound of HFT profit on US stock market equal to 21 billion dollars/year for a 10 second holding period and only 21 million for a 10 millisecond holding period. Duhigg [4] suggested the same 21 billion dollars upper bound, Arnul et al. [9] suggested 1.5 to 3 billion dollars upper bound while Brogaard [2] suggested 3 billion dollars. Baron et al. [6] studied the E-mini SP 500 futures contract from August 2010 to August 2012 and found an estimation of HFT profits equal to 100 million. Aldridge [1] studied the HFT profit on the forex market and concluded that the returns' upper bound is 4 basis point.

As far as we know, there is no equivalent study dealing with recent data on the European Market. In addition, we did not find any paper studying the HFT holding period.

The main goals of this study are to define a theoretical optimal strategy for a HF Trader, to analyze the factors that might explain HFT profit, and to find the optimal holding period according for the bid-ask spread trading cost. This optimal holding period quantifies the low latency advantage effect and helps understanding HF traders' behaviors. The focus of this paper is on aggressive strategies -based on market orders. Limit orders do not increase the adverse selection risk for other participants and are thus widely considered as a harmless activity [5] .

This work is organized as follows: In the first section, some general concepts useful for the rest of the paper are presented. In the second section, the optimal strategy is formulated as a solution of a linear problem. The computation time problem is addressed and some ideas are proposed to enhance the computing performances. In the third section, a one-step omniscient trader method is developed and used to analyze the HFT profit. Results confirm the modest upper bound, discussed above and show a strong dependence of HFT profit on the volatility. Finally, the one-step assumption is relaxed and the methodology, formulated in the second section, is applied to compute the optimal holding period. Results of this section are surprising and show that the optimal trading frequency is not as high as widely assumed.

Notation

Bold, lowercase characters represent vectors, and bold capital characters represent matrices.
In particular, the following denote :

- $\boldsymbol{v}$: A column vector.
- $\boldsymbol{v}^T$: A row vector equal to the transpose of $\boldsymbol{v}$.
- $\boldsymbol{O}$: A matrix which all elements are equal to zero.
- $\boldsymbol{o}$: A vector which all elements are equal to zero.
- $\boldsymbol{I}$: The identity matrix.
- $\boldsymbol{i}$: A vector which all elements are equal to one.
- $\boldsymbol{L}$: A lower full triangular matrix with all non-zero elements equals to one.

1 Preliminaries

1.1 Aggressive HFT

In order to buy/sell a number of shares on an order book driven market, the trader can either match other participants' interests or provide a new offer to the market. For example, **Fig.1** represents an order book with two limits. Some participants are currently willing to buy (Bid side) 100 shares and 70 shares, at 45.5 and 45.4 euros respectively. Other participants are willing to sell (Ask side) 80 shares and 90 shares, at 45.7 and 45.8 euros respectively. At the current state of the order book there are no matching interests. Thus, no transaction is executed.

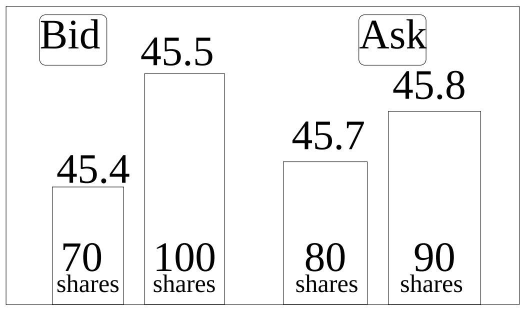


Figure 1: Order book

Suppose a trader wants to buy 50 shares, he can either "hit the order book" and "consume liquidity" by buying 50 shares at 45.7 euros, or post a "buy order" at any price below 45.7 euros. In the first case, the order is called "market order" and the participant is a "liquidity taker". In the second case, the order is called a "limit order" and the participant is a "liquidity provider".

This paper deals exclusively with a liquidity taker trader, i.e. one who uses exclusively market orders. HF traders acting through limit orders can be viewed as liquidity providers to the market, and there seem to be a consensus that providing more liquidity to market participants is harmless, see [12] [10] [8].

This study also focuses on profit made when running a strategy based on short holding periods. Lower frequency strategies are runnable with any framework and thus, are not specific to HFT.

1.2 Data and Framework

This study focuses on the EURO STOXX 50 European liquid stocks. Three years of full daily order book data provided by the "Chair of Quantitative Finance" at Ecole Centrale Paris are used. Snapshots are extracted every 10 millisecond. Auction phases are ignored since traders can not hit the order book during those phases. Thanks to the Mesocentre of the Ecole Centrale Paris, millions of calculations were computed, in a reasonable amount of time, to achieve this study.

2 Omniscient Optimal HFT Strategy

2.1 Problem Formulation

This section aims to mathematically define an optimal strategy relative to some criteria. Knowing the price time series, the available Bid and Ask quantities, and the transaction fees, the following question is answered, "What strategy would have maximized a given utility function?". To achieve this work, the final wealth U_T is considered as the utility function.

A strategy is defined as the vector $\boldsymbol{v}$ such the i^{th} coordinate v_i is the signed number of shares to hold between the time i and the time $i+1$. Given the price time series, $\boldsymbol{p}$, and the chosen strategy, $\boldsymbol{v}$, the final wealth, U_T is to be calculated.

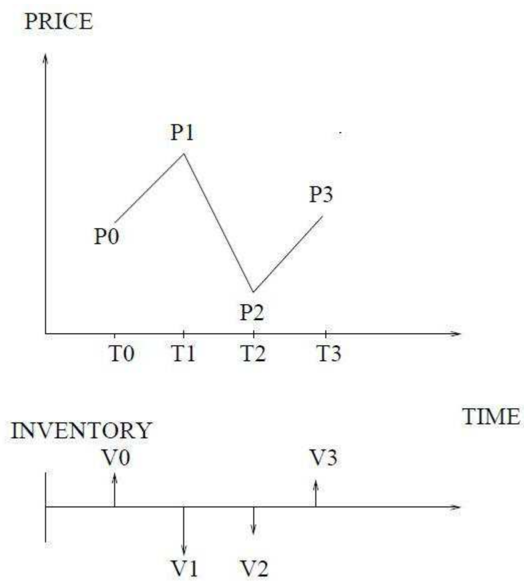


Figure 2: Price and strategy evolution over 4 steps

$\boldsymbol{\delta v}$ is denoted as the variation of $\boldsymbol{v}$ ($\delta v_i = v_i - v_{i-1}$ for $i > 0$). The initial condition $\delta v_0 = v_0$ (before time 0, the folio is empty) is chosen.
Assuming that transaction fees can be assimilated to a proportional cost, λ, $\boldsymbol{U}$ can easily be calculated, for example, at the time 1:

$$U_1 = v_0(p_1 - p_0) - \lambda|v_0|p_0 - \lambda|v_1 - v_0|p_1$$

$$U_1 = -\delta v_0 p_0 - \delta v_1 p_1 + \delta v_1 p_1 + \delta v_0 p_1 - \lambda|\delta v_0 p_0| - \lambda|\delta v_1 p_1|$$

More generally, the wealth U_T obtained by applying a strategy $\boldsymbol{v}$ over T periods is as follows:

$$U_T(\boldsymbol{\delta v}) = \sum_{i=0}^{T} -\delta v_i p_i + p_T \sum_{i=0}^{T} \delta v_i - \lambda \sum_{i=0}^{T} |\delta v_i p_i|$$

Due to the initial condition, a strategy is perfectly defined by giving indifferently $\boldsymbol{v}$ or $\boldsymbol{\delta} v$.

The focus of this study is HFT, thus it is assumed that the portfolio is empty at the end of the period T; $\sum_{i=0}^{T} \delta v_i = 0$. All notations are simplified by dealing only with the best limits of the order book. The general case is detailed in the Appendix. Considering liquidity and trading constraints, the optimal strategy is determined by solving the following problem:
Minimize

$$J_\lambda(\boldsymbol{\delta v}) = \sum_{i=0}^{T} (\delta v_i^+ p_{ask_i} + \delta v_i^- p_{bid_i}) + \lambda \sum_{i=0}^{T} (\delta v_i^+ p_{ask_i} - \delta v_i^- p_{bid_i})$$

Subject to

- $-bidQ_i \leq \delta v_i^- \leq 0$ (Liquidity constraints)
- $0 \leq \delta v_i^+ \leq askQ_i$ (Liquidity constraints)
- $\sum_{i=0}^{T} \delta v_i = 0$ (No overnight position constraint)
- $\delta v_i = \delta v_i^- + \delta v_i^+$ (Definition)
- Min inventory $\leq v_i \leq$ Max inventory (Trading constraints)

2.2 Resolution

Solving the previous optimization problem might seem easy from a mathematical point of view, however, when dealing with high dimensional problems, the simplest linear system might become costly in computation time. This section compares different methods to solve the problem. In particular, the importance of the sparsity when dealing with big data is shown. The key to HFT is to process large amounts of data rapidly. Solving a problem becomes useless if the calculation time is long enough for input data to significantly change. In the next paragraphs, the results obtained using the CVXOPT package and those obtained using the MOSEK solver are compared. For each solver, both dense and sparse formulations of the problem are used.

2.2.1 Framework

2.2.1.1 Sparse Matrices A sparse matrix is a matrix populated mainly by zeros. The fraction of zero elements is called the sparsity of the matrix. In programming, such particularity leads to an important gain of storage space. In stead of storing all the n^2 values of the matrix, only the p non-zero values and their coordinates in the original matrix are stored. Without any loss of the initial information, an important proportion of the storage space is economized. In numerical analysis, most of the powerful solvers correctly handle sparse matrices and take advantage of the sparse structure to economize time when solving numerical problems.

2.2.1.2 CVXOPT Package CVXOPT is a free software package for convex optimization based on the Python programming language. The package provides solvers for linear and quadratic problems. It handles sparse matrices implementations and it is easy to use in any external program.

2.2.1.3 MOSEK Package MOSEK is a large-scale optimization software providing solvers for linear, quadratic, general convex and mixed integer optimization problems. MOSEK handles sparse matrices implementations. The software is not free but provides free academic licenses for research and educational purposes.

2.2.1.4 Matricial Formulation (Dense Formulation) For classic programming languages the problem is described in matricial form as follows:
Minimize

- $\boldsymbol{c}^T \boldsymbol{x}$

Subject to

- $\boldsymbol{G}\boldsymbol{x} \leq \boldsymbol{h}$

Where

- $\boldsymbol{c}^T = \left[p_{ask_0}(1+\lambda), ..., p_{ask_{T-1}}(1+\lambda), p_{bid_0}(1-\lambda), ..., p_{bid_{T-1}}(1-\lambda)\right]$
- $\boldsymbol{x}^T = \left[\boldsymbol{\delta v}^{+T}, \boldsymbol{\delta v}^{-T}\right]$
- $\boldsymbol{G} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{O} \\ -\boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} \\ \boldsymbol{O} & -\boldsymbol{I} \\ \boldsymbol{L} & \boldsymbol{L} \\ -\boldsymbol{L} & -\boldsymbol{L} \end{bmatrix}$
- $\boldsymbol{h}^T = \left[\boldsymbol{askQ}^T, \boldsymbol{o}^T, \boldsymbol{o}^T, \boldsymbol{bidQ}^T, v_{max}, ..., v_{max}, 0, -v_{min}, ..., -v_{min}, 0\right]$

Dimension

- $\boldsymbol{x} \in \mathbb{R}^{2T}$
- $\boldsymbol{G} \in \mathbb{R}^{6T} * \mathbb{R}^{2T}$
- Number of non-zero elements in $\boldsymbol{G}$: $2T^2 + 6T$

2.2.2 Computation times

MOSEK and CVXOPT computation times for several dimensions are compared in **Table 1**.

Dimension T	MOSEK	CVXOPT
100	0.05 s	0.04 s
1000	9.60 s	584.00 s
2000	72.30 s	4156.40 s

Table 1: Computation times

MOSEK is 60 times faster than CVXOPT, however both solvers are so slow compared to the latency needed for a HF strategy.

2.2.3 Variables duplication

Matricial Formulation (Sparse Formulation) In order to reduce the number of non-zero elements, a redundant variable $\boldsymbol{v}$ is introduced. This variable is unnecessary since v_i is perfectly defined knowing $(\delta v_j)_{0 \leq j \leq i}$. The new formulation is:

Minimize

- $\boldsymbol{c}^T \boldsymbol{x}$

Subject to

- $\boldsymbol{G}\boldsymbol{x} \leq \boldsymbol{h}$
- $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{o}$

Where

- $\boldsymbol{c}^T = \left[p_{ask_0}(1+\lambda), ..., p_{ask_{T-1}}(1+\lambda), p_{bid_0}(1-\lambda), ..., p_{bid_{T-1}}(1-\lambda), \boldsymbol{o}\right]$
- $\boldsymbol{x}^T = \left[\boldsymbol{\delta v}^{+T}, \boldsymbol{\delta v}^{-T}, \boldsymbol{v}^T\right]$
- $\boldsymbol{G} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{O} & \boldsymbol{O} \\ -\boldsymbol{I} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{I} \\ \boldsymbol{O} & \boldsymbol{O} & -\boldsymbol{I} \end{bmatrix}$
- $\boldsymbol{h}^T = \left[\boldsymbol{askQ}^T, \boldsymbol{o}^T, \boldsymbol{o}^T, \boldsymbol{bidQ}^T, v_{max}, ..., v_{max}, 0, -v_{min}, ..., -v_{min}, 0\right]$
- $\boldsymbol{A} = [\boldsymbol{I}, \boldsymbol{I}, \boldsymbol{\Lambda}]$
- $\boldsymbol{\Lambda} = \begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 & 0 \\ 0 & \cdots & 0 & 1 & -1 \end{bmatrix}$

Dimension

- $\boldsymbol{x} \in \mathbb{R}^{3T}$
- $\boldsymbol{G} \in \mathbb{R}^{6T} * \mathbb{R}^{3T}$
- $\boldsymbol{A} \in \mathbb{R}^{T} * \mathbb{R}^{3T}$
- Number of non-zero elements in $\boldsymbol{G}$ and $\boldsymbol{A}$: $10T - 1$

Remarks

- In the second formulation, the dimension of the problem is increased by 50%.
- When introducing the redundant variable the number of non-zero elements is reduced from $O(T^2)$ to $O(T)$.

In this paragraph, computation times obtained when solving the new matricial problem are compared to the previous results. For the first formulation, when the computation time is too long (more than 1 hour), it is estimated as $O(T^3)$. Estimated computation time is noted in bold. **Table 2** summarizes the results.

Dimension	MOSEK (Dense)	MOSEK(Sparse)	CVXOPT(Dense)	CVXOPT(Sparse)
100	0.05 s	0.02s	0.40 s	0.04 s
1000	9.60 s	0.07 s	584.00 s	2.80 s
2000	72.30 s	0.12 s	4156.40 s	11.00 s
4000	596.00 s	0.24 s	**33000.00 s**	47.00 s

Table 2: Computation times for dense and sparse formulations

When using the sparse formulation, the computation time decreases spectacularly. In **Fig. 3**, for both formulations, MOSEK is used to compute the solution and computation time is plotted vs problem size (T). It can be concluded, in this case, that rewriting the problem in a sparse form, using a redundant variable, decreases the calculation cost from $O(T^3)$ to $O(T)$.

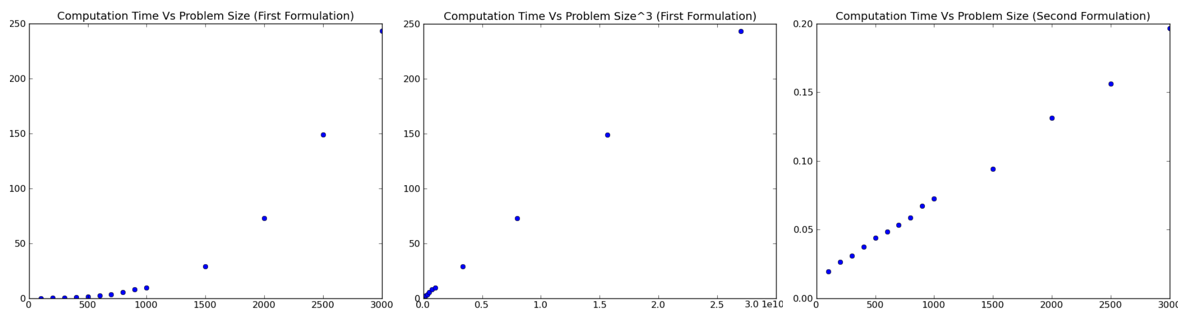


Figure 3: Computation time vs problem size

3 Upper Bound for HFT Strategy and Optimal Holding Period

This section aims to compute an upper bound for HFT profits, to analyze the main factors that explain HFT profitability and to compute an optimal holding period for HF strategy. To this end, an omniscient trader who can observe the future and act accordingly to realize benefits is simulated.

This assumption is not realistic, since the best a trader can do is to predict the future with a small error. However, such results give an idea about the maximum possible HFT profit realized by executing all the profitable trades over 50 stocks for three years.

In the first part of this section, the method presented by Kearns [7] is developed and the HFT profits are explained using different market indicators. In the second part, the one-step method is generalized in the n-steps case using the previous results to compute the optimal strategy and find the HFT optimal holding period.

3.1 Omniscient Order Book Trading - One step

3.1.1 Methodology

The experiment consists of a trader observing the present and the future state of the order book at a given frequency, and taking all profitable positions. Two key time quantities are involved. The first one is the holding period, h, of any taken position. This period has to be long enough for the order book to undergo sufficiently large changes enabling the realization of profits that offset the trading costs due to bid-ask spread crossing, but short enough in order to remain in a high frequency setting.

In fact, a holding period of one millisecond is too short to observe a favorable movement in the order book. A holding period of one minute is too long, and therefore offsets the advantage of rapid exchange access, making the opportunity of profit available to non-high frequency traders.

The second key time quantity is the acting period, m. This quantity is important since it is assumed that the trader does not impact the markets. Indeed, the liquidity taken by the trader when he acts at time t, is returned to the order book when he re-observes it at time $t + m$ to decide to take a new position. It is then clear that a profitable position taken at time t will be available (and then also taken) at time $t + m$ if the order book does not move. This is in accordance with the aim to estimate an upper bound, even if this upper bound can be made arbitrarily high by taking m to be arbitrarily small. Thus, m has to be small enough in order to realize this large bound for the benefits, and large enough in order to avoid the pathological case of taking one profitable position infinitely many times. In addition, to avoid counting artificial profits, the omniscient trader is forbidden from taking positions impossible to be unwind during the next 15 seconds. The order book can show important moves after a "long period" (15 seconds or more) without any change. Thus, the omniscient profitable trade cannot be counted as a HF trade. This step m is chosen to be $m = 10$ milliseconds. This is

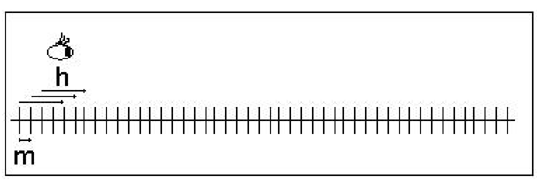


Figure 4: Omniscient Trading : Each m second, the omniscient trader can see the current state of the order book, and its state at the time t+h, he takes thus all possible profitable positions at t and unwind them at t+h.

still very short to have a large overestimation of the profitability, as a winning position can be taken 100 times within a second if the order book does not move enough within that second. This is in accordance with the aim to overestimate the benefits, and avoids the pitfall of very large overestimation.
Another key hypothesis is that the trader is omniscient and thus always makes the good decision.

3.1.2 Results

The different results obtained when running the omniscient strategy over three years of data are analyzed. It came out that HFT profits are modest and negligible compared to traded volumes. It is also shown that profitable trades are very rare for short holding periods.

3.1.2.1 Global Results Here results of running the omniscient strategy over EURO STOXX 50 stocks between 2011 and 2013 are summarized. **Fig. 5** presents the main results of this section. It plots the total profits vs holding periods. The profits decrease rapidly with a decreased holding period. The maximum total profit possible for a holding period of 10 seconds is 85 billion euros and for a holding period of 10 milliseconds is only 4.4 million euros. As discussed in the next paragraph, these sums are modest compared to the traded volume. It can also be noted that the profit in 2011 was significantly higher than 2012 and 2013. This might be explained by a fall in volume and volatility during the last 2 years.

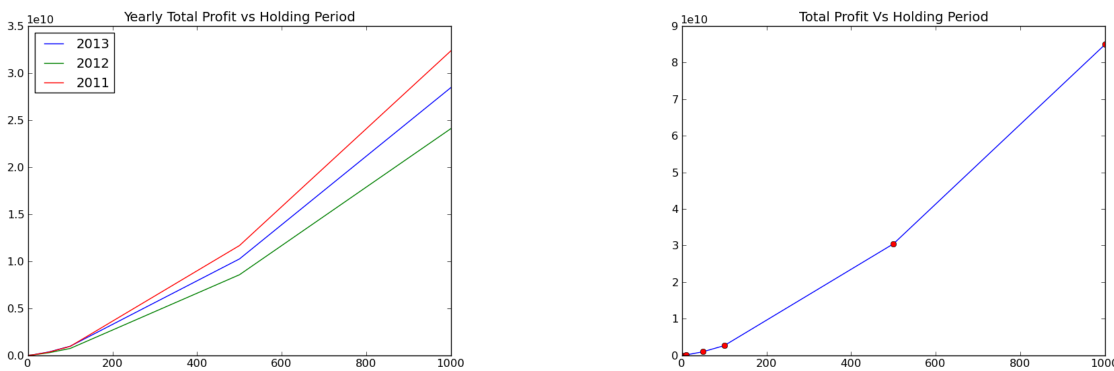


Figure 5: Total profit

To have more familiar numbers, we plot in **Fig. 6** the average profit per stock per day. For a holding period of 10 milliseconds, an omniscient trader, trading aggressively, without transactions fees, taking all profitable decisions at least once, makes on average 136 euros per stock per day! The profit rises up to 2.7 million euros per stock per day for 10 seconds holding period. However, it is impossible to be omniscient for 10 seconds.

Previous results give also an approximation of the possible profit of a non-omniscient trader. Let $U_T(p)$ to be the wealth realized by a trader making predictions with a success probability $p < 100\%$.

A simple approximation gives $U_T(p) = p * U_T(100\%) - (1 - p) * U_T(100\%)$. To verify this formula, a trader with 80% prediction success rate (**Fig. 6**) is simulated. The average profit is approximatevly equal to 60% (A linear regression gives $\beta = 0.599$) of the omniscient average profit, which is coherent with our formula.

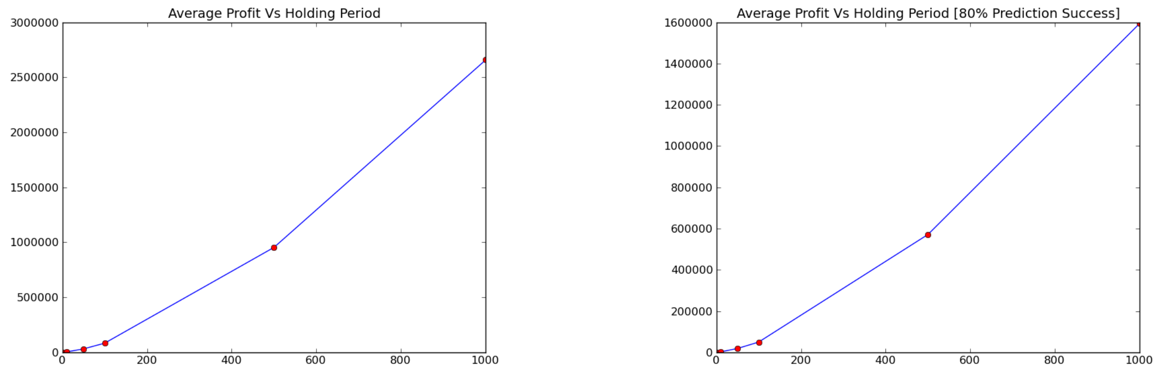


Figure 6: Average profit per stock per day

To understand the causes of small profitability for short holding periods, the average number of trades and the average number of traded shares vs holding period are plotted in (**Fig. 7**). For the 10 milliseconds holding period the average number of trades is 34 and the average number of shares is 27,918. Profitable positions become rare when the holding period is short. This is mainly caused by the bid ask spread that becomes not negligible for small moves of the order book.

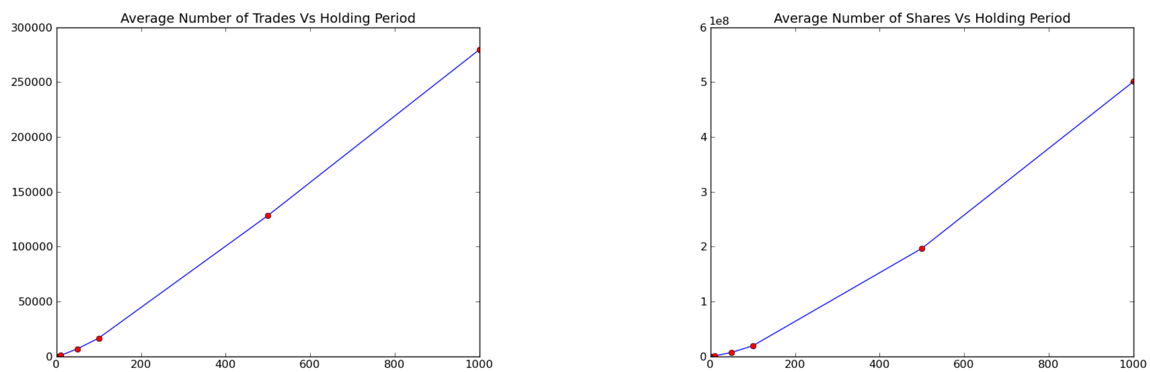


Figure 7: Average trading per stock per day

Besides the fact that profitable positions are rare for short holding periods, **Fig. 8** establishes that they are also less profitable. For the shortest holding period, the average profit by trade is 6.7 euros and the average return is 2.8bps (bps : basis point $= 1\% * 1\%$).

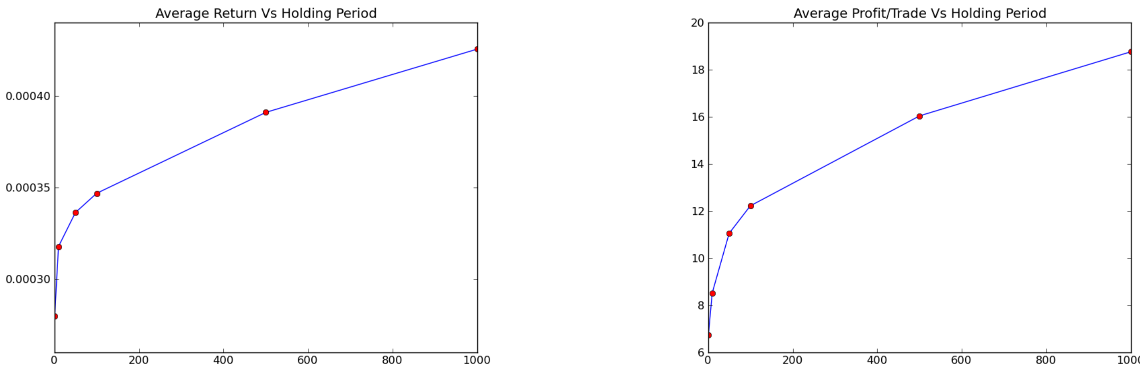


Figure 8: Average profitability

Data used to plot **Fig. 5**, **Fig. 6**, **Fig. 7** and **Fig. 8** are summarized in Table 3

	10 ms	100 ms	500 ms	1 sec	10 sec
Total Profit (2013) [million euros]	1.2	35	359	962	28,468
Total Profit (2012) [million euro]	1.5	30	275	725	24,105
Total Profit (2011) [million euro]	1.6	31	339	947	32,375
Total Profit (All) [million euro]	4.4	97	974	2,634	84,948
Average Profit [euros]	136	3,051	30,562	82,631	2,658,734
Average Number of Trades	34	842	6,873	16,573	279,914
Average Number of Shares	27,918	702,589	7,114,299	19,050,229	501,433,780
Average Return [basis points]	2.8	3.2	3.4	3.5	4.3
Average Profit per Trade [euros]	6.7	8.5	11	12	18

Table 3: Global results

3.1.2.2 Detailed Results In this part the distribution of HFT returns is studied in more detail with the focus on the shortest holding period. **Fig. 9** represents the average daily profit per day during all the studied period, and the density of daily profits. Graphs establish that very profitable days are so rare.

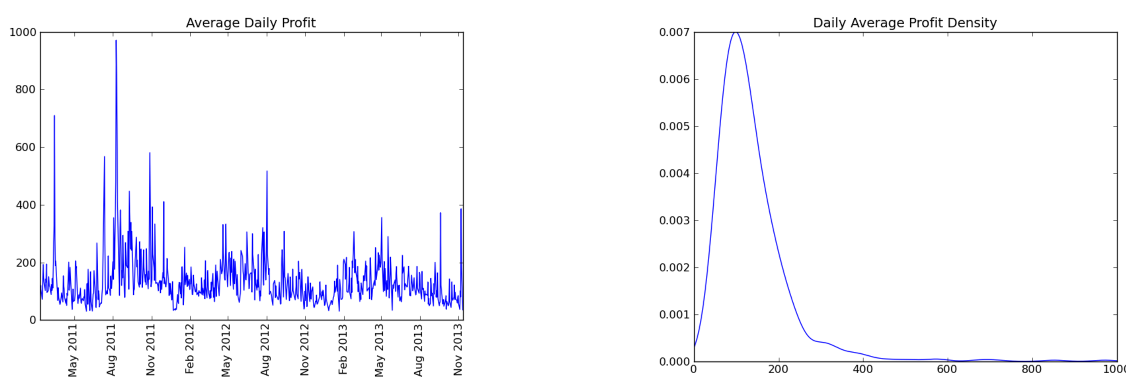


Figure 9: Average profit

In order to understand the main factors driving HFT profits, the daily average profit is plotted vs some features of the EURO STOXX 50. **Fig. 10** examines the relationship between HFT profitability and the Future instrument returns ($\frac{ClosePrice - OpenPrice}{OpenPrice}$). It establishes that profits are more explained by the returns' absolute values than by the returns them selves. A negative correlation (-3%) is observed in the first case, and a positive, more significant, correlation (42%) is observed in the second case. The first result might be explained by the fact that down moves are more brutal (because of agents panic), thus more profitable for HFT traders. The second result is quite intuitive, since an omniscient aggressive trader makes more money when order book shows big moves.

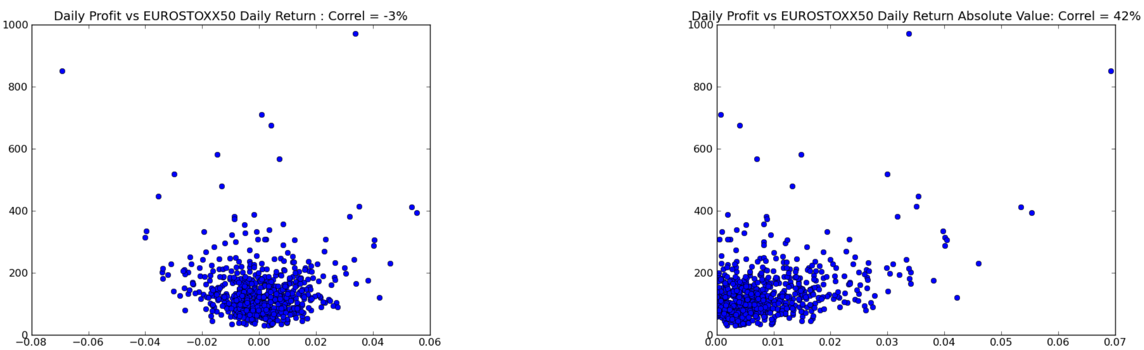


Figure 10: Average profit vs EURO STOXX 50 returns

Since obtained results show that HFT profits are more explained by the volatility than by the return, a better intraday volatility indicator should give results that are more significant. In **Fig. 11** the Daily Range indicator (Daily High - Daily Low) as a proxy of intraday volatility. HFT profits are plotted vs the Daily Range Indicator. The correlation rises up to 64%. In order to keep in mind the relative value of HFT profits, the average daily profit is plotted vs the Future EURO STOXX 50 total traded volume. The correlation is high (56%) which shows that to make more profit, a HFT needs big volumes. Another interesting result is that the best trading day (out of three years) of the omniscient HF aggressive trader (10 milliseconds holding period) ended with less than 50,000 euros of profit. In that same day, 100 billion euros were traded on the Future EURO STOXX 50.

Similar observed effects on temporal analysis are present on cross sectional analysis. HFT performs better on volatile and liquid stocks. In particular, 30% correlation between the stock volatility and the profit made over the stock is observed.

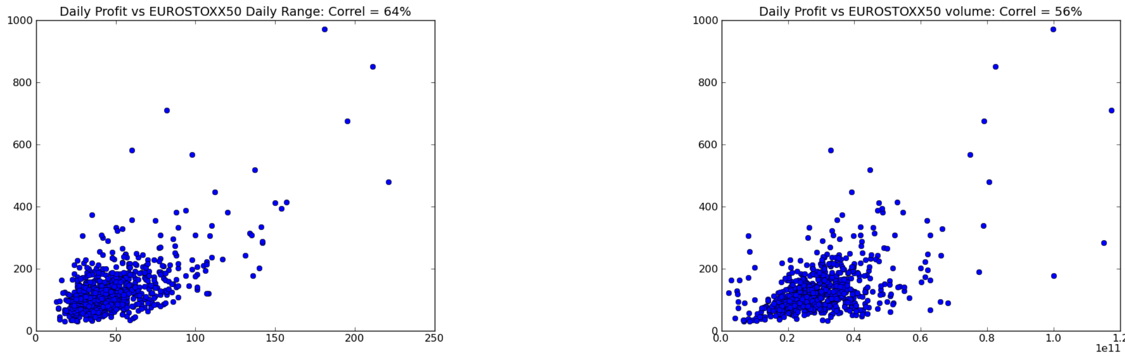


Figure 11: Average profit vs EURO STOXX 50 daily range and traded volume

This section concludes with performances comparisons over the main European markets.

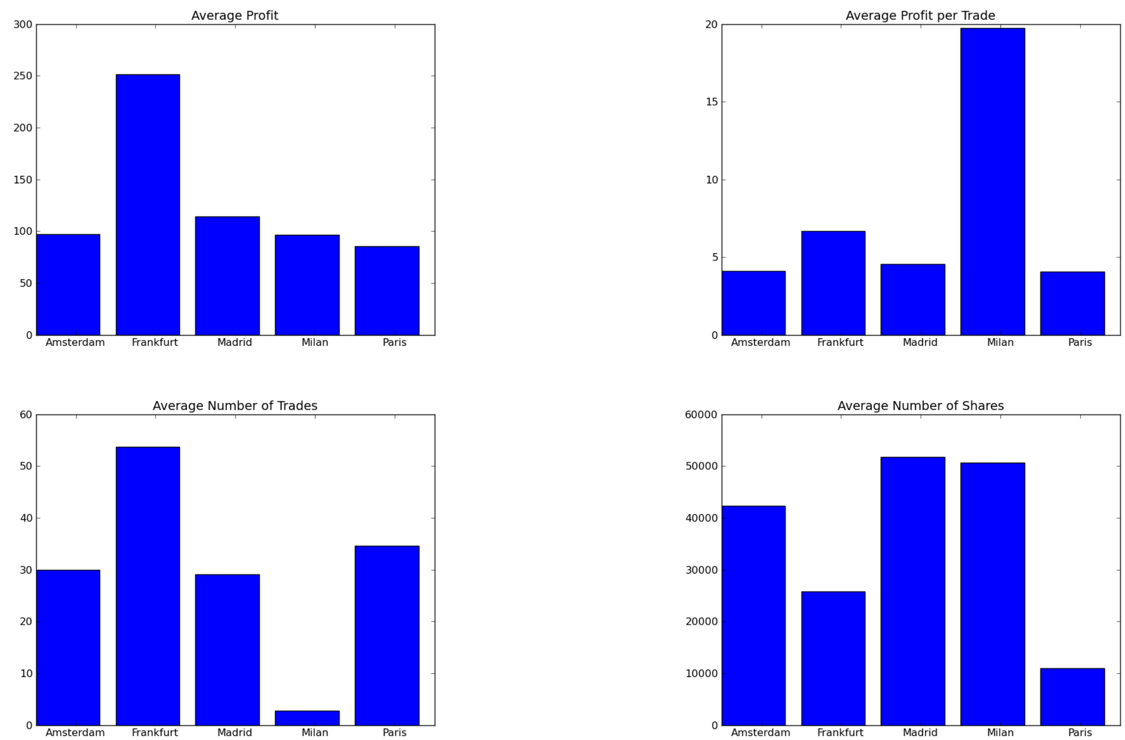


Figure 12: HFT Profitability in main European markets

The graphs establish that in the Italian market, profitable trades are rare. This can be explained by the enormous quantities in the best bid and the best ask. It is rare to observe a big move that consumes all the best limit quantity, however when it happens, the HFT trader can make an important profit per trade due to the big available liquidity. It is also observed that the German market presents more profitable trades due to the big liquidity and small ticks.

3.2 Omniscient Order Book Trading - N steps

In the previous section, empirical results prove that HFT profits are modest for short holding periods. The strategy presented supposes that the trader knows two states of the order book each time; the current state and the next state. The goal of this section is to analyze the optimal strategy in a more general case and to understand the behavior of a trader who can perfectly predict all the changes in the order book during some omniscience period.

3.2.1 Methodology

Similar to the previous section, the experiment consists of a trader observing the present and the future state of the order book at a given frequency, and taking all profitable positions. The new element here is that the trader knows not only the state of the order book at time t and time $t + h$, but also knows all intermediary states. He can switch positions indefinitely under the constraint of having an empty portfolio at the end of each omniscience period. As usual, the trader can buy or sell all the available quantities on the order book without any impact.

The aim of this section is to understand the behavior of a HF trader able to trade at any frequency relative to a 10 milliseconds sampled order book and 10 seconds omniscience period. If low latency advantage were so important, the trader would rapidly switch his positions (each 10 milliseconds in the extreme case). In the other hand, if profit were made on slower moves, the trader would hold his positions for longer periods (10 seconds in the extreme case). For each opened and closed position, the holding period T is computed as the difference between the closing position time and the opening position time. If the trader opens many successive positions without closing the previously opened positions, the assumption is made that positions are closed in the chronological order (first opened, first closed).

Finally, the weighted mean holding period is defined as a weighted (by the quantities) mean of all holding periods. The use of weights is very important; with equal weights, a trader holding 1000 shares for 10 seconds and 1 share for 10 milliseconds, would have a holding period of 5 seconds! For the example of **Fig. 13** the mean holding period is given by $T = \frac{Q_1 T_1 + Q_2 T_2}{Q_1 + Q_2}$.

This measure gives a precise idea about the HFT low latency added value. If HFT traders make the biggest part of their profits on fast trades, the mean holding period should be significantly smaller than the omniscience period.

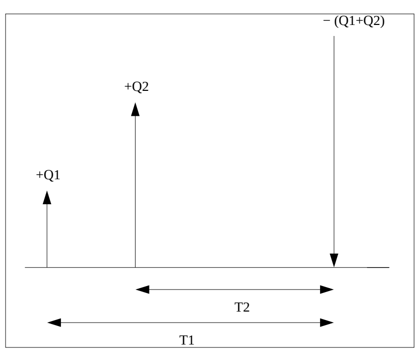


Figure 13: Each position is defined by a quantity and a holding period

3.2.2 Example

To illustrate the methodology, are plotted in **Fig. 14** one stock's mid price evolution over 10 seconds and the corresponding optimal omniscient strategy according for the order book liquidity constraints. In this example, the omniscience period is 10 seconds. The Table 4 shows

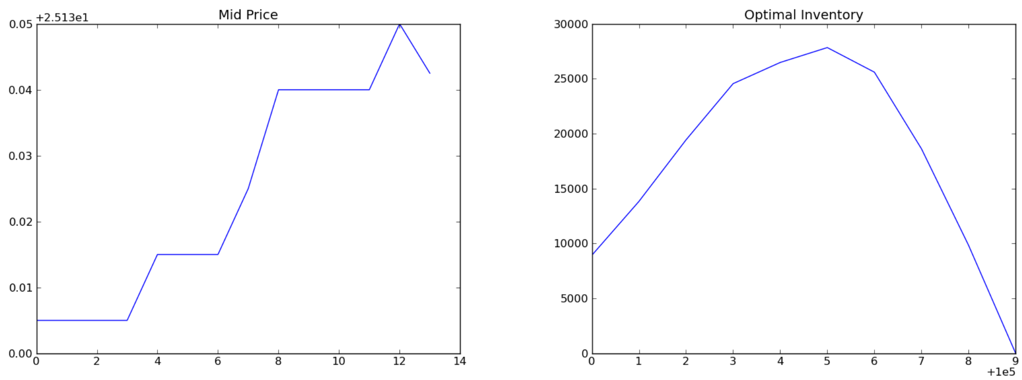


Figure 14: Each position is defined by a quantity and a holding period

the detailed evolution of the trader's portfolio over this 10 seconds period. When a new trade is executed, if the new quantity has the same sign as the existing position, the quantity is added to the list of previous quantities. If the new quantity has an opposite sign, it is used to close the oldest opened position. This rule is used to compute the mean holding period following the formula given in the previous paragraph.

Timer	Trade	Opening Times and Held Quantities	Mean Holding Period (seconds)
00:00	+8928	[00:00] [8928]	-
00:01	+4905	[00:00, 00:01] [8928, 4905]	-
00:02	+5603	[00:00, 00:01, 00:02] [8928, 4905, 5603]	-
00:03	+5121	[00:00, 00:01, 00:02, 00:03] [8928, 4905, 5603, 5121]	-
00:04	+1927	[00:00, 00:01, 00:02, 00:03, 00:04] [8928, 4905, 5603, 5121, 1927]	-
00:05	+1357	[00:00, 00:01, 00:02, 00:03, 00:04, 00:05] [8928, 4905, 5603, 5121, 1927, 1357]	-
00:06	-2239	[00:00, 00:01, 00:02, 00:03, 00:04, 00:05] [6689, 4905, 5603, 5121, 1927, 1357]	6.00
00:07	-6980	[00:01, 00:02, 00:03, 00:04, 00:05] [4614, 5603, 5121, 1927, 1357]	6.73
00:08	-8786	[00:02, 00:03, 00:04, 00:05] [1431, 5121, 1927, 1357]	6.63
00:09	-9836	[] []	6.29

Table 4: Portfolio evolution and mean weighted holding period computation

3.2.3 Results

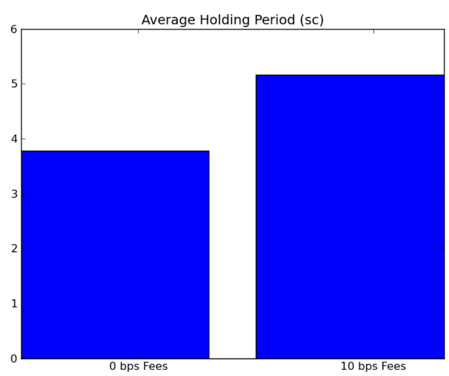
The first graph of **Fig. 15** shows the main results of this section. A trader who knows perfectly all the order book evolution for 10 seconds with 10 milliseconds sampling, and trades with 0 costs, would have an average a holding period of 3.8 seconds. This holding period is 380 times greater than the smallest possible holding period ; 10 milliseconds.

Such result mitigates the claim that low latency advantage is the main key of HFT profits. Making money when hitting the order book and paying the bid ask cross, is very diffcult.

When the trader is subject to 10 bps trading costs, the holding period increases to 5.1 seconds. The number of trades decreases from 106,000 trades to only 10,000 trades per stock per day.

In the second graph of **Fig.15** the holding period is plotted vs the bid ask spread. It can be seen that the holding period depends strongly on trading fees. When trading becomes costly, only very profitable trades are executed. Those trades should provide a return higher than the fees. Such high returns are more likely observed on long holding periods.

The dependence of the holding period on Bid Ask spread is less clear. However, a positive correlation of 17% can be seen. The Bid Ask spread represents the average crossing cost. A positive correlation is consistent with the fact that holding periods increase with trading costs.



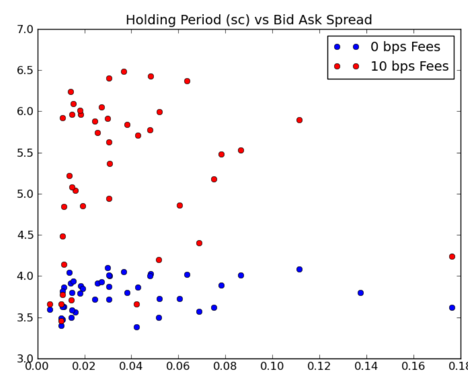


Figure 15: Average holding period

4 Conclusions

This paper provides a large empirical study dealing with EURO STOXX 50 stocks over the last three years. To compute an objective upper bound of aggressive HFT profits, a one-step omniscient strategy is applied. The results confirm other papers' studies dealing with other markets (Forex, US Equities..) [7] [4] [9] [2] [6] [1]. Profits are rather modest and even negligible for the shortest holding periods.

To get rid of the fixed holding period hypothesis, a new method to compute an optimal HFT strategy is introduced, the n-steps omniscient strategy. This method is used to compute a new measure; the weighted mean holding period. Results show that this period is 400 times greater than the smallest possible period. In other words, an omniscient trader is trading on average with a frequency 400 times slower than the highest available frequency, which shows that hitting the order book rapidly in order to take advantage of low latency information asymmetry is not that profitable.

We acknowledge stimulating discussions with Sebastien LEFORT, Axel BREUER and Riadh ZAATOUR.

References

[1] I. Aldridge. How profitable are high-frequency strategies? Oct 2009.

[2] J. Brogaard. High frequency trading and its impact on market quality. 2010.

[3] B. Appelbaum D. Cho. Aobama's 'volcker rule' shifts power away from geithner. *Washington Post Staff Writer*, Jan 2010.

[4] C. Duhigg. Stock traders find speed pays, in milliseconds. *The New York Times*, Jul 2009.

[5] R. Riordan J. Brogaard, T. Hendershott. High frequency trading and price discovery. 2013.

[6] A. Kirilenko M. Baron, J. Brogaard. The trading profits of high frequency traders. 2012.

[7] Y. Nevmyvaka M. Kearns, A. Kulesza. Empirical limitations on high frequency trading profitability. 2010.

[8] Albert J. Menkveld. High frequency trading and the new market makers. 2011.

[9] J. Saluzzi. S. Arnuk. Latency arbitrage: The real power behind predatory high frequency trading. 2009.

[10] A. Menkveld T. Hendershott, C. Jones. Does algorithmic trading improve liquidity? *The Journal of Finance*, 2011.

[11] J. Tobin. A proposal for international monetary reform. *Eastern Economic Journal*, 1978.

[12] S. Zwick. High-frequency trading: Good, bad or just different? 2010.

Appendix: Multi Limits Case Formulation

The optimal strategy problem formulated in 2.1 can be generalized in the multi limits case. K is denoted as the number of limits available and x_i^j as the value of x relative to the limit j at the time i.
Minimize

$$J_\lambda(\boldsymbol{\delta v}) = \sum_{i=0}^{T}\sum_{j=0}^{K-1}(\delta v_i^{j+} p_{ask_i{}^j} + \delta v_i^{j-} p_{bid_i{}^j}) + \lambda \sum_{i=0}^{T}\sum_{j=0}^{K-1}(\delta v_i^{j+} p_{ask_i{}^j} - \delta v_i^{j-} p_{bid_i{}^j})$$

Subject to

- $-bidQ_i{}^j \leq \delta v_i^{j-} \leq 0$ (For each j - Liquidity constraints)
- $0 \leq \delta v_i^{j+} \leq askQ_i{}^j$ (Liquidity constraints)
- $\sum_{i=0}^{T-1} \delta v_i = 0$ (No overnight position constraint)
- $\delta v_i^{+} = \sum_{j=0}^{K-1} \delta v_i^{j+}$ (Definition)
- $\delta v_i^{-} = \sum_{j=0}^{K-1} \delta v_i^{j-}$ (Definition)
- Min inventory $\leq v_i \leq$ Max inventory (Trading constraints)