HIGH FREQUENCY TRADING AND LEARNING

JASMINA ARIFOVIC[1], CARL CHIARELLA[2], XUE-ZHONG HE[2] AND LIJIAN WEI[3,*]

[1] Simon Fraser University, Department of Economics
Burnaby, BC V5A 1S6, Canada

[2] University of Technology Sydney, Business School
PO Box 123, Broadway, NSW 2007, Australia

[3] Sun Yat-Sen University, Sun Yat-Sen Business School
No.135, West Xingang Road, Guangzhou 510275, China

Date: First Version: June 2015; this version: January 22, 2016.

Acknowledgement: We thank participants in the 2014 Sydney Economics and Financial Market Workshop and the seminars at the University of Technology Sydney, Sun Yat-Sen University and Tianjin University for valuable comments. We also thank Shu-Hen Chen, Giulia Iori, Blake LeBaron, Fabrizi Lillo and Paolo Pellizzari for valuable comments. Financial support from the Australian Research Council (ARC) under the Discovery Grants (DP110104487, DP130103210) and the National Natural Science Foundation of China (NSFC) Grants (71320107003) is gratefully acknowledged.

**Corresponding author.*

Emails: `arifovic@sfu.ca (Arifovic)`, `Carl.Chiarella@uts.edu.au (Chiarella)`, `Tony.He1@uts.edu.au(He)`, `weilijian@gmail.com (Wei)`.

Electronic copy available at: http://ssrn.com/abstract=2771153

Abstract. This paper introduces a limit order market model of fast and slow traders with learning to examine the effect of high frequency trading (HFT) and learning on limit order markets. We demonstrate that informed HFT makes significant profit from trading with other traders and, more importantly, it is the learning and information advantage that plays more important role than the trading speed in generating HFT profit. Overall HFT increases market liquidity consumption and supply, trading volumes, bid-ask spread, volatility and order cancelations, reduces order book depth, improves information dissemination efficiency, and generates significant event clustering effect in order flows. Interestingly, the speed of HFT is positively related to trading volume and spread and negatively related to market depth; however it has an U-shaped relation to liquidity supply and market efficiency, but an invert U-shaped relation to HFT profit and liquidity consumption. The findings provide some insight on the profitability of HFT and the current debates and puzzles about HFT.

Key words: High frequency trading, learning, informed traders, limit order market, genetic algorithm, liquidity

JEL Classification: G14, C63, D82

Electronic copy available at: http://ssrn.com/abstract=2771153

1. Introduction

High frequency trading (HFT) is becoming a dominant trading technology in financial markets. However there are some on-going questions on how high frequency (HF) traders use and learn from market information. There are also and some debates and puzzles about the impact of HFT on financial markets, including price discovery, information efficiency, volatility, order flow and liquidity. To address these questions and provide better understanding of the empirical puzzles, we introduce a limit order market model of fast and slow traders, use a genetic algorithm with a classifier system to model sophisticated learning from information available in the market, and examine the effect of learning and HFT on limit order markets.

This paper contributes to the literature in three aspects. The *first* is on how HFT's learning affects the information processing and trading profit. We show that the sophisticated learning based on genetic algorithm with a classifier system makes HF traders use more market information than low frequency (LF) informed and uninformed traders, in particular the information related to fundamental value, moving-averages and depth imbalance of buy and sell sides. This effect becomes more significant as the trading speed increases. Also, due to their private information about the fundamental value and high trading speed, informed HF traders make significant profit from trading with LF traders. The profit is driven by the high trading speed, but more importantly, information advantage and learning. In fact we show that HFT's trading speed has an invert U-shaped relation to HFT profit, meaning that there is a trade-off between trading speed and profit of HFT. In addition, HFT reduces the profit opportunity for the informed LF traders significantly and the loss for uninformed LF traders, meaning that uninformed LF traders can actually benefit from HFT through the learning from more frequently released information to the market. Furthermore HFT's profit improves with less competition among informed HF traders and high volatility in fundamental value.

The *second* contribution is on how HFT affects order submission behaviour of fast and slow traders and the overall market liquidity. We show that HFT increases trading volume and order cancelation. It also improves both liquidity consumption and supply in the overall market, although it affects order submission differently for different types' traders. In particular, HFT makes informed HF traders and LF (informed and uninformed) traders submit less market orders but more aggressive limit orders. Due to high trading volume and low execution of limit orders, both liquidity supply and consumption are improved. This effect becomes more significant with learning and more competition among informed HF traders. However with high volatility in fundamental value, HFT leads to an increase in both market and aggressive limit orders. In addition the trading speed has a nonlinear effect on the order submission and liquidity supply and consumption, displaying a seemingly U-shaped relation to limit order submission and hence to liquidity supply, but a significant invert U-shaped relation to market order submission and hence to liquidity consumption. Furthermore HFT generates significant event clustering effect in order flows characterized by the positive serial correlations of market orders and all types of limit orders. This effect is mainly driven by high trading speed of HFT, not affected by learning, fundamental volatility, and information lag of uninformed traders.

The *third* contribution focusing on the impact of HFT on market efficiency, volatility and spread in the limit order market. We show that HFT increases trading volumes, bid-ask spread, volatility and order cancelations, reduces order book depth, improves information dissemination efficiency and hence price discovery. We also find that market efficiency reduces and volatility increases with an increase in the volatility of the fundamental value and a decrease in the competition among informed HF traders. In addition, we find that the speed of HFT is positively related to trading volume and spread but negatively related to market depth, but interestingly, it has an U-shaped relation to market efficiency. Overall, we provide a broad framework to better understand the learning and trading activities of HFT, and its impact on limit order markets. The results lead to some implications on market policy and design, provide some insight on the profitability of HFT and the current debates and puzzles about HFT.

This paper is closed related to the literature on sophisticated trading algorithm and profitability of HFT. HFT is dominating the markets not only due to its trading speed advantage but also due to its sophisticated trading algorithm and learning from microstructure information, such as short-term price movements, order flows, order books and market conditions. It is well recognized that HFT is very profitable and the current HFT literature focuses more on trading speed (Hoffmann (2014)) rather than strategic trading behavior or both. As pointed out by Easley, de Prado and OHara (2013) : *"HFT is here to stay. The current speed advantage will gradually disappear, as it did in previous technological revolutions. But HFT's strategic trading behavior... is more robust."* Therefore understanding of HFT's strategic trading behaviour is very important. In this paper, we model the sophisticated learning and trading decision using genetic algorithm (GA) with a classifier system as a learning mechanism for traders to learn from microstructure and historical information and to update trading rules.

Since introduced firstly by Holland (1975), GA has been used to examine learning and evolution in economics and finance (Arifovic (1994, 1996) Arthur et al., (1997) and Routledge (1999, 2001)). GA learning is a search heuristic based on historical performance mimicing the evolutionary process of natural selection including selection, mutation, and crossover. The use of the classifier system in GA was firstly introduced by Holland (1975) and then in economic and financial market models.[1] A trading rule generated by GA contains two components; market condition and trading action. The classifier system is used to classify the market information or condition for traders and help them to process various information and submit orders accordingly. A trading action corresponds to order types, market or limit orders, and the aggressiveness of limit orders. The advantage of the GA with a classifier system is that traders can learn to trade from very high dimensional state space in limit order books. More recently, GA with classifier system has been used to study the learning and order submission in limit order markets.[2] In particular, without HF traders, Chiarella, He and Wei (2015) study the evolution of trading rules and the effect of learning on limit order markets.

[1]See, for example, Marimon, McGrattan and Sargent (1990), Allen and Carroll (2001), Lettau and Uhlig (1999), and SFI-ASM models including Arthur, Holland, LeBaron, Palmer and Tayler (1997) and LeBaron, Arthur and Palmer (1999).

[2]For example, Wei, Zhang, He and Zhang (2014) allow uninformed traders to use GA to learn from limit order market prices and find the GA learning improves information efficiency.

This paper examines, under HFT's learning, what types of information traders use when updating their trading rules and what drives HFT's profit. The limit order market model of fast and slow traders introduced in this paper is based on Chiarella et al. (2015) and the difference is to allow a fraction of traders to be HF traders so that we can study the impact of HFT. Without HFT, Chiarella et al. (2015) find that, measured by the average usage of different group of market information, trading strategies under the learning become stationary in the long run. Also the average information usage frequency for uninformed traders is higher than for informed traders, though informed traders pay more attention to the last transaction sign while uninformed traders pay more attention to technical rules. By allowing a fraction of informed traders to be HF traders, this paper shows that HFT makes HF traders use more market information than LF traders, in particular the information related to fundamental value, moving-averages and depth imbalance of buy and sell sides. Therefore HFT affects how traders process information when updating their trading rules. Also, the trading speed is positively correlated to the information usage.

With information asymmetry, learning and HFT, our model allows us to examine the driving sources of HFT's profit. Consistent with empirical literature, we first confirm that informed HF traders make significant profit and reduce the profit of informed LF traders. We also find that HFT helps uninformed LF traders reduce their loss, meaning that uninformed LF traders can benefit from HFT. Our results show that learning makes HFT more profitable, which is consistent with Chiarella et al. (2015), however, uninformed HFT is not profitable. This implies that HFT's profit is more driven by information advantage and learning, while the trading speed has an invert U-shaped relation to HFT's profit. Intuitively, HFT helps traders exploring profit opportunity quickly and reducing pick-off risk; however, a very high trading speed increases the competition among informed HF traders, which reduces their profit opportunity. Furthermore an increase in competition among informed HFT reduces their profit. These results imply that there should be some optimal trade-off between trading speed and the number of informed HF traders that maximizes HFT's profit.

There are some debates on how HFT affects market liquidity. In general, HFT can be classified (see Martinez and Rosu (2011)) to either passive HFT, such as market-marking that is likely to use limit order to provide liquidity and manage inventories, or active HFT that uses sophisticated algorithm, including statistical arbitrage and machine learning (Kearns and Nevmyvaka (2013)), to exploit various sources of information from fundamentals and order flows.[3] Passive HFT has different impact on the market from active HFT. With passive market-making, HFT narrows the spread and increases the depth, therefore increases market liquidity. However, Kirilenko and Lo (2013) argue that *"...contrast to a number of public claims, high frequency traders do not as a rule engage in the provision of liquidity like traditional market makers. In fact, those that do not provide liquidity are the most profitable and their profits increase with the degree of 'aggressive' liquidity-taking activity."* Therefore examining the effect of active HFT on liquidity is critical to understand these debates. This paper focuses on active and informed HFT, instead of market-making. Given the complexity of HFT, there is very limited models on active HFT,[4] The model introduced in this paper allow us to address the impact of active HFT. The result on high profit from informed HF traders with learning provides a supporting evidence to the argument of Kirilenko and Lo (2013).

There are also some debates on how HFT affects price discovery, market volatility, bid-ask spread and order book depth. [5] Some empirical studies[6] show that at least

[3]There are other categories for HFT. For example, according to SEC (2010), there are four types including passive market marking, arbitrage, structural, and directional HFT. Essentially, the first type belongs to passive HFT and the other three types are active HFT.

[4]The informed HFT models such as Martinez and Rosu (2011) and Biais, Foucault and Moinas (2011) are not limit order models. The HFT model in limit order markets of Hoffmann (2014) is a three-periods model, which is not capable of capturing some impact on market quality and order flow dynamics. Though there are some agent-based models on the market-making HFT and flash crash without learning (see, for example, Vuorenmaa and Wang (2014)).

[5]Zhang (2010) and Hasbrouck (2013) find that HFT does not help to improve price discovery but increases volatility; however, Brogaard (2012) and Brogaard, Hendershott, and Riordan (2014) find that HFT improves price discovery and may reduce intraday volatility.

[6]See Brogaard et al. (2014), Kirilenko, Samadi, Kyle and Tuzun (2011), Hendershott, Jones and Menkveld (2011).

some active HFT traders have superior information than others. Empirical studies also show that informed HFT improves information efficiency or price discovery (see Brogaard et al. (2014), Hendershott et al. (2011), Martinez and Rosu (2011) and Biais et al. (2011)), increases volatility (Martinez and Rosu (2011), Kirilenko et al. (2011) and Biais et al. (2011)) and trading volume (Martinez and Rosu (2011)). It also increases the adverse selection of LF traders (Biais et al. (2011)). In this paper, we focus on informed and active HFT. Our analysis on the profitability of informed and uninformed HF traders and the impact of informed HFT on market efficiency, volatility, and trading volume provide consistent results to these empirical findings. Our result shows however a different impact of learning on market efficiency with and without HFT. Without HFT, Chiarella et al. (2015) show that market information efficiency is improved when uninformed traders learn, but not necessarily when informed traders learn. With HFT, we find that market information efficiency is always improved.

Most of the empirical studies are based on the aggregate data and it is difficult to identify the effect of different types of traders under active HFT. The model introduced in this paper allows us to examine not only the aggregate market impact, but also individual order submission behaviour. We show that HFT not only increases trading volume and order cancelation, but also improves market liquidity consumption and supply. On order submission, because of the information and speed advantage, HFT makes informed HF traders and LF traders submit less market orders but more aggressive limit orders. This is different from Chiarella et al. (2015) who show that, without HFT, learning makes uninformed traders submit less aggressive limit orders and more market orders. We also find that HFT increases trading volumes, bid-ask spread, volatility and order cancelations, reduces order book depth, improves information dissemination efficiency and hence price discovery. These findings are consistent with Brogaard (2010) who finds that HF traders actually supply less order depth than other traders and Kim and Murphy (2013) who find that market spreads were much worse than have been reported in the U.S. markets, but inconsistent with Gai, Yao and Ye (2012) who point out that in aggregate level, the bid-ask spread and order book depth do not change much.

The rest of this paper is organized as follows. The model is introduced in Section 2. Section 3 examines evolutionary dynamics of the GA. Section 4 examines the effect of HFT on order profit, order flow, market liquidity, information efficiency, and volatility. Section 5 concludes. The details of the design of trading rules and the GA learning mechanism are presented in the Appendices.

2. Model

We consider a limit order market with asymmetric information and HFT. As in Chiarella et al. (2015), traders are either informed or uninformed and their trading rules are generated and updated endogenously through the GA learning based on their private and public information. Different from Chiarella et al. (2015), we allow a fraction of traders to become HF traders who enter the market more frequently nd fast.[7]

2.1. **The limit order market.** There are N risk neutral traders who trade a risky asset in a limit order market. To accommodate both low and high frequency trading, we let time period t, defined by $(t-1, t]$, be a short-time interval, such as milliseconds or seconds, and time period T, defined by $(T-1, T]$, be a long-time interval, such as minutes or hours. Typically, $T = mt$ for some positive integer m. We assume that HF traders enter the market at the short-time interval t and LF traders enter the market at the long-time interval T. For example, if we allow traders to enter the market every period and set t, say 10 seconds, as one period and $m = 6$, then HF traders enter the market at every 10 second interval, while LF traders enter the market at every one minute interval.

The fundamental value v_t of the risky asset at short-time period t follows a random walk process with an initial fundamental value of v_o. Innovations in the fundamental value v_t occur according to a Poisson process with parameter ϕ. If an innovation occurs, the fundamental value either increases or decreases with equal probability by κ tick sizes. In a benchmark case, among the traders, there are N_H HF informed

[7]Following the literature, the speed in the HFT defined in this paper is "chronological time", rather than the "volume clock", which is essentially the core of HFT according to Easley et al. (2013). We leave the HFT under the volume clock for the future study.

traders,[8] N_I LF informed traders and N_U LF uninformed traders with $N_H < N_I < N_U$ and $N_H + N_I + N_U = N$. The informed HF traders know the fundamental value of the current time period v_t when they enter the market in the short-time period t. The informed LF traders only know the (same) fundamental value of the time period v_T when they enter the market in the long-time period T.[9] This setup allows HF traders to react to the news and trade much faster than LF traders. The information is short-lived, meaning that uninformed traders know the fundamental values with a time lag of $\tau > 0$ measured in units of the long-time period T, which is also called the information lag or information-lived time. For example, if T is one minute and $\tau = 360$, then the LF uninformed traders know the fundamental value lagged by one trading day[10]. The asymmetric information structure between informed and uninformed traders and the short-lived information for uninformed traders are similar to Goettler, Parlour and Rajan (2009). When entering the market, traders submit orders to buy or sell at most one unit of the asset. Transactions take place based on the standard price and time priorities in limit order markets. We let $p_t = p_{t-1}$ if there is no transaction between time $t-1$ and t, and p_T be the last transaction price over the long-term period T.

When a trader enters the market at time t', he observes a number of pieces of common information from the market price and the limit order book, including the

[8]We also consider the case in which some informed and uninformed traders are HF traders in this paper. When the HF traders are the uninformed only, our results show that the HF uninformed traders do not benefit from HF trading and short-lived information, making significant loss; therefore we do not consider this special case.

[9]This is a very important feature of informed HFT traders. For example, when a news announces, the HFT can react to the news and submit orders directly to the order book of the exchange via the direct market access (DMA), which is much faster than a broker system used by slow traders. For example, if t is a millisecond and m is 60,000, then T is one minute. When a fundamental innovation occurs in the 59,900 millisecond and no innovation occurs from 59,901 millisecond to 60,000 millisecond, the HFT traders who enter the market at 59,900 millisecond know the fundamental value $v_{59,900}$, while slow informed traders who enter the market at the 60,000 millisecond only know the same fundamental value by lagged 100 milliseconds.

[10]There are 360 minutes of one trading day in Australia stock markets.

current transaction price $p_{t'}$, the current bid $b_{t'}$ and ask $a_{t'}$ prices, the mid-price (bid-ask midpoint) $p^m_{t'} = (a_{t'} + b_{t'})/2$, the current bid-ask spread $s_{t'} = a_{t'} - b_{t'}$, the depth of the limit order book, the depth at the bid $d^b_{t'}$ and the ask $d^a_{t'}$, the depth of the buy side $d^{buy}_{t'}$ and the sell side $d^{sell}_{t'}$, the buy or sell initiated transaction sign $p^{\pm}_{t'}$ (+ for a buy and - for a sell).[11] For HF traders, they observe the market price p_t at short-time period t; while for LF traders, they observe the market price p_T at long-time period T. All the traders observe the average market price $\bar{p}_{T,\tau} = [p_{T-1} + p_{T-2} + \cdots + p_{T-\tau}]/\tau$ over the last τ long-time periods,[12] For a LF trader, when he enters the market at the long-time interval of T, he observes the current order book information and the historical prices at the long-time interval of T. Each of type i (HF informed, LF informed and uninformed) traders enter the market, submit either market orders to trade or limit orders then exit and reenter the market after some periods according to a Poission process with parameter λ_i. Upon reentry, the trader cancels his previous limit order and submits a new order.[13]

2.2. **Trading rules and GA learning.** When a trader enters the market, he uses a GA with a classifier system to process the market information and chooses the best trading rules to buy or sell one share with either a market order or limit order (including aggressive limit order, limit order at the quote, or unaggressive limit order away the quote). The details of the GA learning is given in the appendix. The difference between the informed (both LF and HF) and (LF) uninformed traders is that the trading decision to buy or sell is determined by the private information of the fundamental value for the informed traders, while it is part of learning for the uninformed traders.

[11]In some extreme case with very small probability, there is no bid when the buy limit order book is empty or no ask when the sell limit order book is empty. In these cases we let $a_{t'} = 1.01p_{t'-1}$ when the sell limit order book is empty and $b_{t'} = 0.99p_{t'-1}$ when the buy limit order book is empty.

[12]To examine the learning effect, we keep the same market information structure on the average price between the LF and HF traders and record the last trading prices over the long~~-term~~ time period T.

[13]Otherwise, to reduce pick-off risk, the unexecuted limit order expires after DT periods.

Table 1. The groups of the classifier rule (CRs): Group FV based on CR1 about the expected fundamental value; Group TR based on CR2 to CR4 about the technical rules; Group QS based on CR5 to CR7 about the change in quotes and bid-ask spread; Group DI based on CR8 and CR9 on order book depth imbalance; and Group TS based on CR10 on the last transaction sign. Here, $v_t^i = v_t^H = v_t$ for informed HF traders, $v_t^i = v_t^I = v_T$ for LF informed traders, $v_t^i = v_t^U = v_{T-\tau}$ for LF uninformed traders, $\bar{p}_{t,\tau} = \bar{p}_{T,\tau}$ and $\bar{p}_{t,\tau/2} = \bar{p}_{T,\tau/2}$.

Group	Num	CR	Description
FV	CR1	$p_{t'}^m > v_t^i$	The mid-price is higher than the expected fundamental value.
TR	CR2	$\bar{p}_{t,\tau} > v_t^i$	The average market price of last τ periods is higher than the expected fundament value.
	CR3	$p_{t'}^m > \bar{p}_{t,\tau}$	The mid-price is higher than the average market price of last τ periods.
	CR4	$\bar{p}_{t,\tau/2} > \bar{p}_{t,\tau}$	The average market price of last $\tau/2$ periods is higher than the average market price of last τ.
QS	CR5	$s_{t'} > s_{t'-1}$	The current spread is larger than the last spread.
	CR6	$a_{t'} > a_{t'-1}$	The current ask is higher than the last ask.
	CR7	$b_{t'} > b_{t'-1}$	The current bid is higher than the last bid.
DI	CR8	$d_{t'}^a > d_{t'}^b$	The current depth of the ask is larger than the current depth at the bid.
	CR9	$d_{t'}^{sell} > d_{t'}^{buy}$	The current depth of the sell side is larger than the current depth of the buy side.
TS	CR10	$p_{t'-1}^{\pm}$	Last transaction sign (last market order is buy or sell).

A trading rule has two parts: the market condition part and the trading action part, as in Chiarella et al. (2015). The market condition part describes the current market conditions, such as "the current spread is larger than the last spread". Motivated by empirical studies, we consider ten of the most important market conditions, corresponding to ten classifier rule listed in Table 1. Under the classifier system, we use a binary string with 10 bits to classify all the market conditions.

If a classified rule is "true", the corresponding bit value is "1"; otherwise it is "0". If a trader does not consider a classified rule when making a trading decision, the corresponding value of the classified rule is labeled as"#". The action part contains the buy-sell decision and order type, such as aggressive limit buy. We use binary strings with 3 bits to indicate all the actions and list them in Table 2. For simplicity, the aggressive limit order is one-tick price inside the bid-ask spread than the best quote while unaggressive limit order is one-tick price away from the best quote. We then combine these two parts together to generate trading rules. To start with, each trader has a number of trading rules. When a trader enters the market, he chooses the best trading rules whose condition parts match the current market conditions.

Table 2. The actions or order types

Action (buy)	Binary code	Description
MB	000	Market buy
ALB	001	Aggressive limit buy
LBA	010	Limit buy at the bid
ULB	011	Unaggressive limit buy
Action(sell)	Binary code	Description
MS	111	Market sell
ALS	110	Aggressive limit sell
LSA	101	Limit sell at the ask
ULS	100	Unaggressive limit sell

Each trader uses GA to update the trading rules. The learning mechanism of GA is an evolutionary process based on the principles of natural selection, crossover and mutation. Selection means that a trading rule is selected by a tournament mechanism based on its performance. Crossover means that a trader chooses two trading rules with high performances as parents, splits each trading rule into two parts at a random bit and then swaps the two parts to create two new trading rules as children. Mutation means that a trader selects a trading rule with high performance as a parent and makes a random bit change of the parent trading rule to a different value to create a new trading rule. GA then uses these new children (trading rules) to replace a fraction of trading rules with weak performance. GA updating takes place every fixed number of long-time periods called one generation. Initially, traders randomly choose order types so the condition parts of the trading rules contain trinary string with many "#". During the evolution process, more market conditions are used and hence the bit of the corresponding market condition changes to "1" or "0". By counting the number of "#" bits for each of the ten market conditions in trader's trading rules, we are able to measure the usage of each of the classifier rule. By examining the evolutionary dynamics of the average usage of each of the market conditions within each group of traders in each generation, we can measure how different types of traders use the GA to process the information via the classifier rule. We refer readers to Chiarella et al. (2015) or Appendix A for more details of the design of trading rules and the learning mechanism of the GA.

2.3. **Experiment design.** The total population of traders is $N = 1000$. To examine different aspect of HFT and the effect of trading speed, learning, the number of informed HF traders and fundamental volatility, we consider 11 cases listed in Table 3.[14] To examine the effect of HFT, we first consider a case, denoted by NH, of no HF traders and all traders are LF traders with the GA learning, which is the benchmark case in Chiarella et al. (2015). Based on some empirical studies on the probability of informed trading (PIN) (see, for example, Yan and Zhang (2012)), we choose the proportion of the informed traders to be 10% in the NH case, which corresponds to 100 informed traders and 900 uninformed traders, as in the benchmark case of Chiarella et al. (2015). To examine the impact of HFT, we next consider a benchmark case, denoted by HF, by allowing 10% of informed traders in case NH to become HF traders; therefore there are 10 informed HF traders, 90 LF informed traders and 900 LF uninformed traders in case HF. Intuitively, the number of HF

[14]In fact, there are three elements of HFT that affect the market, including trading speed, private information, and learning. To examine the private information effect, we also consider a case in which there are some uninformed HF traders and all the informed traders are LF traders. The results show that those uninformed HF traders make a significant loss. Hence we do not include this case and report the results in the paper. Since our model is a zero-sum game, uninformed traders always lose to informed traders, the uninformed traders may trade for external motivation, such as private value.

traders can affect the trading and competition among traders. To examine the effect of the number of informed HF traders and the competition among informed HF traders, we consider two cases, denoted by 5H and 20H, for 5 and 20 informed HF traders, respectively. In addition, to examine the effect of uninformed HFT traders, we also consider a case, denoted by IUH, with 5 informed and 5 uninformed HFT traders in case NH. To analyze the effect of learning, we then consider a case, denoted by NL, in which the informed HF traders in case HF do not learn (while the LF traders still learn) and choose order type randomly.[15]

Table 3. Experiment design. Here IH, UH, IL and UL are the numbers of HF informed and uninformed traders and LF informed and uninformed traders, respectively, λ_H is the arriving rate of HF traders in the HF-time, κ is the volatility of the fundamental value, and τ is the information lag.

Case	Description	IH	UH	IL	UL	λ_H	κ	τ
HF	Benchmark	10	0	90	900	1/6	4	360
NH	without HF	**0**	0	100	900	1/6	4	360
NL	HF **without learning**	10	0	90	900	1/6	4	360
5H	less HF traders	**5**	0	95	900	1/6	4	360
20H	more HF traders	**20**	0	80	900	1/6	4	360
3λ	more frequency	10	0	90	900	**1/3**	4	360
12λ	less frequency	10	0	90	900	**1/12**	4	360
48λ	much less frequency	10	0	90	900	**1/48**	4	360
LV	lower volatility	10	0	90	900	1/6	**2**	360
SL	shorter lag	10	0	90	900	1/6	4	**180**
IUH	uninformed HF	**5**	**5**	90	900	1/6	4	360

For the parameters in the NH case, let the initial fundamental value $v_0 = \$20$, market price $p_0 = v_o = \$20$, and the tick size is \$0.01. The innovation process and volatility are given by a Poisson process with a rate of $\phi = \frac{1}{6}$ and $\kappa = 4$,

[15]In this case, the informed traders know the buy/sell decision, they choose each of the four types of orders with equal probability.

as in Chiarella et al. (2015). This means that, on average, the innovation of the fundamental value occurs once every 6 HF time periods (one minute) and each innovation changes the fundamental value by 4 ticks (either increasing or decreasing by 4 ticks with equal probability). For the information lag, let $\tau = 360$ LF time periods which means that the information-lag of the fundamental values for the uninformed traders is one trading day that assumed to be six hours, which is also the maximum order survival time ($D = 360$). Let HF time period as one period in the simulation time, and $m = 6$. This implies, for example, if one HF period is 10 seconds, then a LF time period is one minute in real markets.[16] To examine the effect of fundamental volatility, we consider a case denoted by LV with a lower volatility $\kappa = 2$. To examine the effect of information lag, we also consider a case denoted by SL with shorter $\tau = 180$.

Traders can enter, reenter the market, and revise the previous limit orders. In the benchmark case HF, the LF traders follow a Poisson process with arrival rate of $\lambda_L = 1/60$ in the LF time period, while the arriving rate for the HF traders is $\lambda_H = \frac{1}{6}$ in the HF time period. With one HF period of 10 seconds and one LF time period iof one minute, this means that each LF trader enters the market once per hour and each HF trader enters the market once per minute on average. One way to measure the trading speed of HFT is how often traders can enter the market to trade. To analyze the effect of trading speed, we consider three cases, denoted by 3λ, 12λ, and 48λ, with the HF traders' arriving rate $\lambda_H = 1/3, 1/12$ and $1/48$, respectively, meaning that each HF trader enters the market once per 30 seconds, 2 minutes, and 8 minutes on average, respectively. Therefore, the HF traders are the fastest in case 3λ and the slowest in case 48λ.

For the evolution process of GA, let $\beta = 0.2$ be the discount rate of historical performance, the crossover rate be 0.1 and the mutation rate be 0.3, as in Chiarella et al. (2015). For HF traders, the evolution process is active on average of 60 HF

[16]More realistically, HFT may enter the market in milliseconds or even microseconds. To simplify the analysis and speed up the simulations, we choose $m = 6$ so that HF traders enter the market six times faster than LF traders.

time periods, which means that on average one generation of is 10 minutes; for LF traders, one generation is 360 LF time periods (one trading day).

For statistical significance, we run 30 simulations. Since the GA needs sufficient learning time to obtain optimal trading rules, each simulation runs 432,000 HF time periods (or 72,000 LF time periods). Then the evolution process is active 7,200 times (generations) for the HF traders and 2,000 times (generations) for the LF traders.

3. Evolutionary dynamics of the GA

In this section, we first study the evolutionary dynamics of the GA for all types of traders in the benchmark case HF. We show that, measured by the usage frequency of the classified rule groups, the evolution of the GA becomes stationary in the long run in this case. We then examine the effect of HFT on the information usage among different types of traders.

3.1. **The evolutionary dynamics of the GA.**

We first examine the evolutionary dynamics of the GA under the HFT. As in Chiarella et al. (2015), we use the usage frequency (probability) γ_j^i of classified rule group $j(j = FV, TR, QS, DI, TS)$ of type $i(i = IH, IL, UL)$ traders to examine the evolutionary dynamics of the GA. The GA becomes stationary if the mean of γ_j^i becomes stationary in the long run. To calculate γ_j^i, when a trader selects a trading rule to trade, we count the number of "#" bits in the corresponding classifier rule. For example, if the condition part of a trading rule is "##1#1 00#1#", the bits for the classifier rule CR1, CR2, CR4, CR8 and CR10 are "#", these CRs are not counted; while the bits for CR3, CR5, CR6, CR7 and CR9 are counted. In this way, we calculate the total counts for each classified rule group for all the traders from the same type. Then the usage frequency of a classified rule group of each type traders is calculated by the ratio of the total counts of the classified rule group used to the total trading times of all the traders from the same type during one generation. For example, when all the informed HF traders trades for 1,000 times and the total usage of CR1 is 500 during one generation, then the usage frequency of CR1 for the informed HF traders is 0.5 for the generation. Put differently, when an informed HF trader trades, the usage probability of CR1 for each trading is 0.5 on average during the generation.

Different types information are characterized by different classified rule groups. We calculate the average usage frequency of each classified rule group. For example, the usage frequency are 0.50, 0.52, 0.56 and 0.57 for CR1, CR2, CR3 and CR4, so the $\gamma^i_{FV} = \gamma^i_1 = 0.5$, and $\gamma^i_{TR} = \frac{1}{3}(\gamma^i_2 + \gamma^i_3 + \gamma^i_4) = \frac{1}{3}(0.52 + 0.56 + 0.57) = 0.55$. Consequently, we calculate the average usage frequencies of all the classified rule groups for each type of traders over different generation and examine the evolution of the information usage. The results are reported in Figure 1.

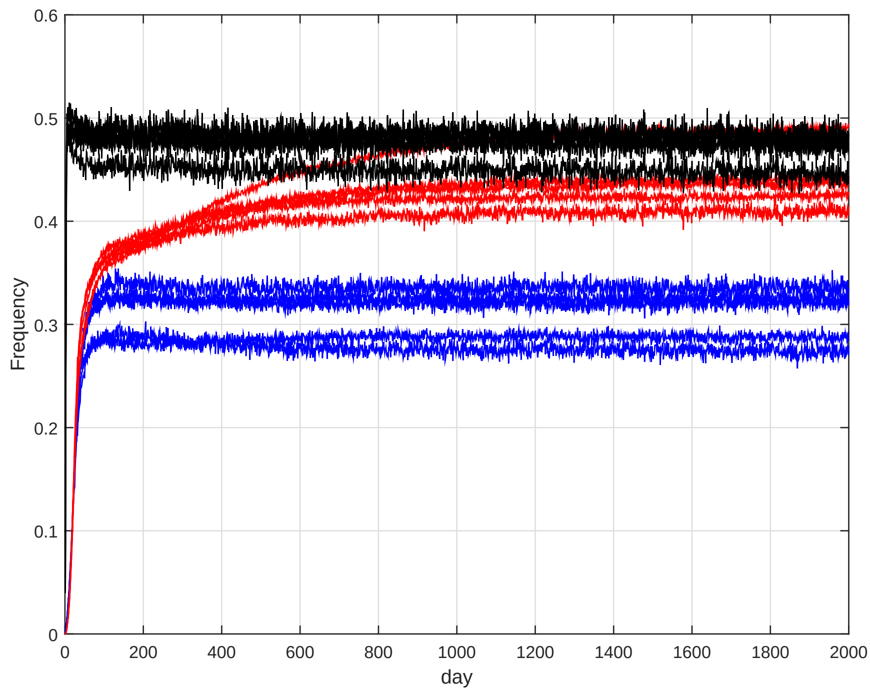


FIGURE 1. The evolutionary dynamic of the classified rule groups of informed HF traders and LF informed and uninformed traders under the GA.

Figure 1 shows that the usage frequencies of all the CR groups grow quickly in the early generations and then settle down and fluctuate around some mean levels in the long run. A statistic test reported in Appendix B shows that all the usage frequencies for all types of traders become stationary after the initial 1800 generations.[17] The speed of becoming stationary is very fast for the informed HF traders (about 50

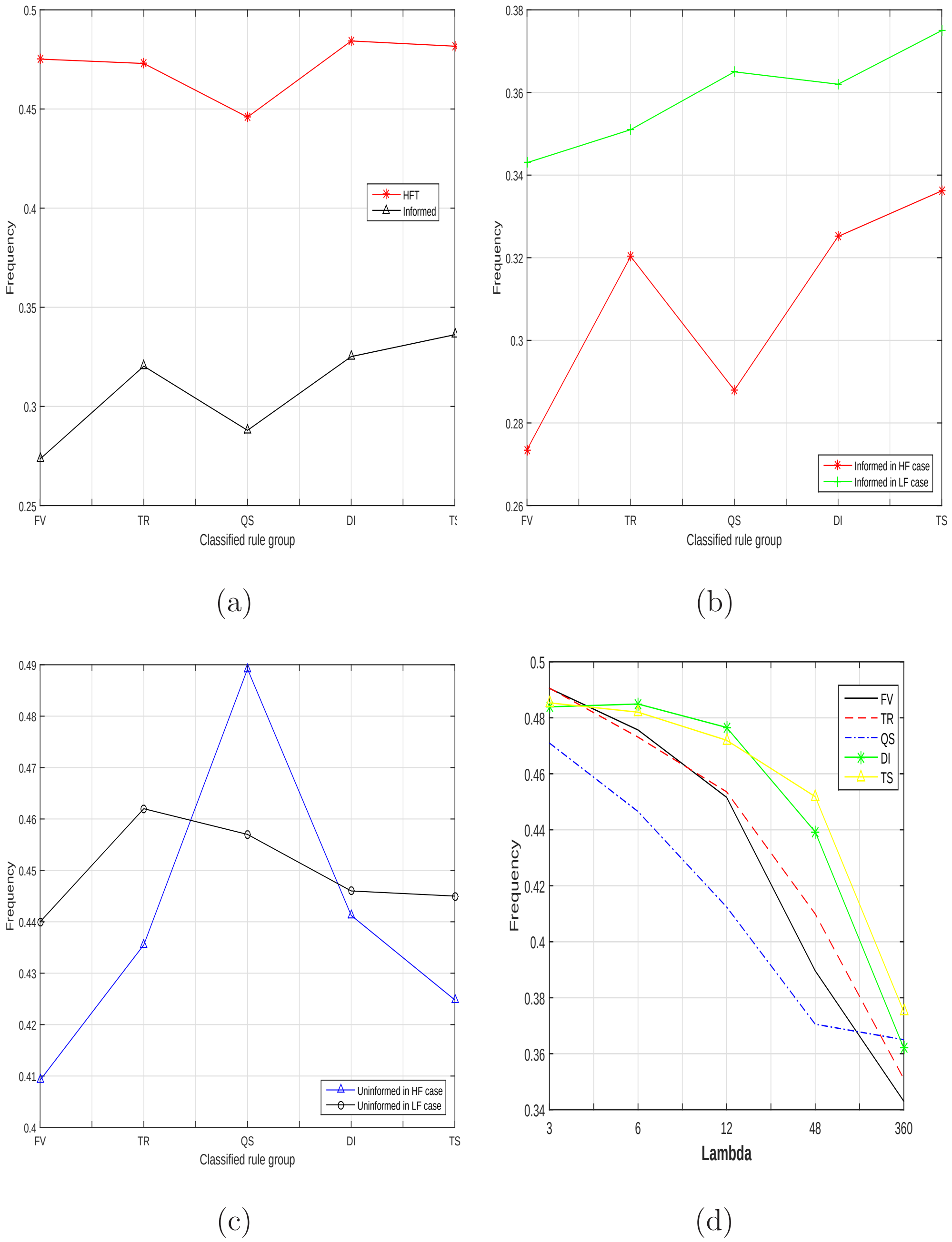
[17]We use the last 200 generation data to examine the stationarity of the usage frequencies. We divide the data into two equal parts and using the ANOVA test. Tables 11, 12 and 13 in Appendix B show that the mean, maximum and minimum value of γ^i_j are very close between the two parts,

generations), followed by the LF informed traders (about 100 generations), and then by the LF uninformed traders. Intuitively, with the same learning mechanism, the learning of the informed HF traders becomes more effective when they enter the market and trade more often. For the uninformed LF traders, the evolution is rather slow (about 1800 generations). One possible explanation is that, when the informed traders trade more often, they release more information to the market; however, because of the slow learning process, the uninformed traders face more adverse selection risk and the released information from the informed HF traders becomes more noisy. As we discuss in the next section, learning may also help informed HF traders to manipulate the order book and thus reduce information efficiency. Therefore HFT makes the evolution process of LF uninformed traders rather slower. However the LF uninformed traders are able to learn eventually. The analysis in the next section is based on the last 200 generations during which the learning processes of all types of traders become stationary.

3.2. **Information usage among different types of traders.** Given the stationarity of the usage frequencies over the last 200 generations, we now examine how different types of traders use the information differently. Intuitively, due to different trading speed and asymmetric information, different traders use the information differently. The trading speed also affects how traders use the information. We report the mean values of the usage frequencies for all five CR groups for different types of traders and for different trading frequency of informed HF traders over the last 200 generations in Figure 2. Based on Figure 2, we have obtained the following results.

First, the usage of all the information is significantly higher for the informed HF traders than for the LF informed traders, illustrated in Figure 2 (a) that plots the usage patterns among the CR groups for HF and LF informed traders in the HF case. It shows that the usage of all the market information for the informed HF traders (about 0.4-0.5) is significantly higher than that for the LF informed traders (about 0.25-0.35) across all the CR groups. This finding is consistent with Brogaard

with very small standard deviations. The p-values of more than 0.05 in ANOVA imply that trading rules of all types of traders are stationary over the last 200 generations.



(a) (b) (c) (d)

FIGURE 2. The patterns of γ_j^H among CR groups for (a) HF and LF informed traders; (b) LF informed traders with and without HFT; (c) LF uninformed traders with and without HFT; together with (d) the usage frequency among CR groups with different trading frequencies λ_H of HF traders

et al. (2014) who find that HFT uses more prices and order book information. Also, compared to the LF informed traders, the informed HF traders use relatively more information on the fundamental value (FV), which is consistent with Martinez and Rosu (2011) who find that HF traders react faster to news.

Secondly, HFT make the LF traders, both informed and uninformed, use less information on all the classified rule groups, except the quotes and bid-ask spreads (QS) for the uninformed LF traders, as illustrated in Figure 2 (b) for the LF informed and (c) for the LF uninformed traders. Because of the endogenously evolutionary process of the GA, the HFT also affects how market information is used by LF traders. For the LF informed traders, we compare their information usage in Figure 2 (b) with and without HFT. Compared to the NH case (of LF traders), HFT makes the LF informed traders use less information on all classified rule groups, in particular on the fundamental value (FV) and the quotes and bid-ask spread (QS). This reduction on the information usage is mainly due to the speed and information advantage of the informed HF traders who release the information to the market more quickly. This in turn reduces the information advantage and learning efficiency of the LF informed traders. As we show later that the informed HF traders reduce the profit of LF informed traders, which makes them have less incentive to learning actively, resulting the reduction in the information usage. For the LF uninformed traders, because the informed HF traders release more information on the quotes and bid-ask spreads, which become more valuable for the uninformed traders, leading to an increase in their usage of information related to group QS.

Finally, the information usage frequency of the HF traders increases with the trading speed. Figure 2 (c) shows that when the trading speed, measured by the Poisson arriving rate λ_H increases from $\lambda_H = 1/360$ to 1/48, 1/12, 1/6 and then 1/3, the HF traders enter the market faster and the usage frequencies are also increasing.[18] But the trend slows down when the trading speed is very high (for $\lambda_H = 1/6$ and 1/3), in particular for the CR groups DI and TS.

Overall, we find that the informed HF traders use more information for trading than the LF informed traders, and their usage frequencies increase with the speed of the HFT. With the HFT, the LF traders also use less information in general, except

[18]We consider that the NH case as a special case that the HF traders has the same trading frequency with LF informed traders.

the information on the quotes and bid-ask spread. These changes in the information usage under the HFT affect the trading behaviour of all traders and therefore the market, we examined next.

4. The effect of HFT

In this section, we first examine the effect of HFT on order profit and order submission for different types of traders. We then examine the impact of HFT on market, including market information efficiency, volatility, and liquidity. The analysis is based on the last 36,000 LF time periods (the last 100 generations of LF traders) when the information usages for all types of traders are stationary.

4.1. **Order profit.** It is commonly believed that HFT is more profitable than LFT, however it is not clear whether the profit is more driven by private information, learning or the speed. To understand different sources of HFT profit, we consider all the cases in Table 3 and report the profit results in Tables 4 and 5. The order profit is defined by $r_t = v_t - p_t$ for executed buy orders and $r_t = p_t - v_t$ for executed sell orders. The order profit per trade r is in ticks, while the total profit R and the average profits $\bar{R}$ from each group over the 100 generations are in dollars. The results lead to a number of observations.

First, informed HF traders profit the most so the speed matters. In the HF case, the order profit per trade, the total and average profits are higher for the informed HF traders than the LF informed traders. In particular, their total and average profits are much higher (about 15 times on the total and average profits), though the order profit per order is much smaller, comparing to case NH without HFT. The result is consistent with the common observation in financial markets that though HFT makes small profit per trade, it gains significant aggregate returns.[19] However, the result in Table 5 shows that such profit opportunity disappears for the HF uninformed traders. For case IUH with 5 informed and 5 uninformed HF traders, together with 90 informed and 900 uninformed LF traders, Table 5 shows that the

[19]According to a report from Barron's in 2010, for example, Renaissance Technology's Medallion fund (a quantitative HFT fund) returned 39% in 2009 and has a 3 year annual compound return of 62.8%.

Table 4. Order profit, here r_i is the order profit per trade in ticks, R_i is the total order profit and $\bar{R}_i$ is the average profit in dollars for group i with $i = IH, IL, UL$ for the HF informed, LF informed and LF uninformed traders, respectively.

Case	r_{IH}	r_{IL}	r_{UL}	R_{IH}	R_{IL}	R_{UL}	$\bar{R}_{IH}$	$\bar{R}_{IL}$	$\bar{R}_{UL}$
HF	2.33	2.22	-1.23	3,162	601	-3,763	316.20	6.68	-4.18
NH	N/A	35.19	-3.51	N/A	10,684	-10,684	N/A	106.84	-11.87
NL	0.67	2.53	-0.52	941	689	-1,630	94.10	7.66	-1.81
5H	5.72	4.85	-1.84	4,316	1,354	-5,670	863.2	14.25	-6.3
20H	0.50	1.23	-0.52	1,307	299	-1,606	65.35	3.74	-1.78
3λ	0.30	1.88	-0.36	586	503	-1,089	58.60	5.59	-1.21
12λ	5.54	4.60	-1.87	4,610	1,212	-5,822	461.00	13.47	-6.47
48λ	16.04	15.18	-2.56	4,155	3,969	-8,124	415.50	44.10	-9.03
LV	1.04	0.91	-0.51	2,559	587	-3,146	255.90	6.52	-3.50
SL	2.32	2.23	-1.21	3,130	605	-3,736	313.30	6.72	-4.15

Table 5. Order profit for case IUH with 5 informed and 5 uninformed HF traders, together with 90 informed and 900 uninformed LF traders, here r_i is the order profit per trade in ticks, R_i is the total order profit and $\bar{R}_i$ is the average profit in dollars for group i with $i = IH, UH, IL, UL$ for the HF informed, the HF uninformed, LF informed and LF uninformed traders.

r_{IH}	r_{UH}	r_{IL}	r_{UL}
7.29	-6.43	1.50	-3.32
R_{IH}	R_{UH}	R_{IL}	R_{UL}
5,804	-2,920	1,722	-4606
$\bar{R}_{IH}$	$\bar{R}_{UH}$	$\bar{R}_{IL}$	$\bar{R}_{UL}$
1160.80	-584.00	19.14	-5.12

informed HF traders make even more significant profit, while the uninformed HF traders make a significant loss, comparing to the LF traders. This indicates that both the speed and private information are necessary to generate profit opportunity for HFT.

Secondly, learning helps the informed HF traders to make significant profits. Comparing the profit results for cases HF and NL in Table 4, the only difference is the learning of the informed HF traders. Without learning, the average order profit per order drops (from 2.33 to 0.67 ticks), and the total and average profits also drop significantly (by about 3 times) for the informed HF traders. Notice that the profits (losses) for the LF informed (uninformed) traders are reduced (increased). When there is no HFT, Chiarella et al. (2015) also show that learning helps informed traders to make profit. However, if we compare the average profit $\bar{R}$ between cases HF and NH, the informed HF traders make about 3 times profit to the LF informed traders. Because of the information advantage, the informed traders know the buy/sell decision but learn the order types (either market or limit orders). Therefore the speeding learning reduces the pick-off risk for the informed HF traders, which improves their profit significantly.

Thirdly, there is a trade-off between the trading speed and profit opportunity and a high trading speed does not necessarily generate high profit. Comparing cases 48λ, 12λ, HF($\lambda_H = 1/6$) and 3λ, the results in Table 4 show that the order profit per order decreases as the trading speed of the informed HF traders increases. However, the total and average profits increases initially (for 48λ and 12λ) and then decreases (for HF($\lambda_H = 1/6$) and 3λ). This implies that there is an optimal trading speed and a trade-off between speed and profit for HFT. Intuitively, the informed traders make their profit by trading with LF traders, in particular the LF uninformed traders. When informed HF traders trade too often, they compete among themselves, which reduces their reading opportunity with LF traders, and hence the profit opportunity.

Fourthly, HFT reduces the profit opportunity for the LF informed traders and the loss for the LF uninformed traders, though the informed traders make profits and uninformed traders make losses in general. Comparing the order profits of LF traders in cases NH and HF of Table 4. The order profit per trade for the LF informed traders drops significantly (from 35.19 ticks in the NH case to 2.22 ticks in the HF case), while the loss of the uninformed LF traders reduces (from 3.51

ticks to 1.23 ticks per trade). The total and average profits are reduced significantly for the LF informed traders (by about 50 times), while the total and average losses are also significantly reduced (by about 2.8 times) for the LF uninformed traders. This implies that, with the same learning mechanism, the LF informed traders loose their information advantage to the informed HF traders. However, the LF uninformed traders benefit significantly from more frequent information release from the informed HF traders.[20]

Furthermore, the competition of informed HF traders reduce their order profit significantly. When the number of informed HF traders is reduced from 10 in the HF case to 5 in the 5H case, the competition among the informed HF traders is reduced, which improves the profit for the HFT. The results in Table 4 show that the average order profit is more than doubled for the informed HF traders while the average order profit for the LF traders is also improved. However, when the number of informed HF traders is increase from 10 in the HF case to 20 in the 20H case, the order profit per trade is reduced to less than 1 tick for the informed HF traders, even smaller than the order profit of LF informed traders. The total order profit and average profit of the informed HF traders are also reduced significantly. The result is consistent with the industry report that the profit of HFT is shrinking much after 2013 due to competition of HFT.[21]

Finally, the HFT becomes more profitable for the informed, in particular the informed HF traders as the fundamental volatility increases. By comparing the profit of the informed HF traders between cased HF and LV (with high and low volatility respectively), the order profit per order for the informed HF traders is

[20]This result implies that, without considering the transaction cost, there is a strong incentive for LF informed traders to become HF traders. However, as more informed traders become HF traders, their profits are reduced due to competition. Therefore there should be an equilibrium threshold number of the informed HF traders at which informed traders are indifferent between LF and HF trading. Below the threshold, HFT is more attractive; above the threshold, LFT is more preferred. This conjecture is partially verified in the following by examining the effect of the number of informed HF traders.

[21]See the report by Greg MacSweeney on Wall Street & Technology, http://www.wallstreetandtech.com/trading-technology/hft-profits-shrinking-the-data-doesnt-lie/a/d-id/1267900.

more than doubled and the total profit increases by 1.23 times as the fundamental volatility is doubled. Therefore information becomes more valuable for the HFT when the information uncertainty increases.

In summary, the results provide supporting evidence to our intuition that it is the combination of information, learning and trading speed that makes the HFT more profitable. More importantly, we find that information and learning play a very important role, while the trading speed has an invert U-shape relation to the HF trading profit. Therefore when the speed gradually disappears, the learning and strategic trading become critical for HFT, as highlighted by Easley et al. (2013). In addition, the profit improves for the informed HF traders with high fundamental volatility and less competition among themselves. Also the LF informed traders loss heavily from the HFT while the LF uninformed traders actually benefit from the HFT.

4.2. **Order choice, liquidity supply and consumption.** We now examine how the HFT and learning affect the order choice, liquidity supply and consumption of different types of traders. Based on the order types in Table 2, we introduce four types of aggregate orders according to the order aggressiveness: $MO = MB + MS$ the aggregate market buy and sell orders; $ALO = ALB + ALS$ the aggregated aggressive limit buy and sell orders; $LOA = LBA + LSA$ the aggregate limit buy and sell orders at the best quotes; and $ULO = ULB + ULS$ the aggregative less aggressive limit buy and sell orders. To better understand the effect on order choice and aggregate orders submitted by HF and LF traders, for each type of traders, we use the fractions of each types of aggregate orders in the total orders submitted to measure the order submission behaviour. We also use the submission rate, taking rate, and execution rate to measure liquidity supply and consumption. The *submission rate (SR)* is the ratio of the number of the limit orders to the total number of orders, measuring the liquidity supply. The *taking rate (TR)* is the ratio of the number of market orders to the number of the executed orders, measuring the liquidity consumption. The *execution rate (ER)* is the ratio of the number of executive limit orders to the number of the total submitted limit orders, measuring order execution (or cancellation). For all the cases in Table 3, we report the results in Table 6. For each case, the middle column reports the percentages of four types of orders for each type of traders and the aggregate percentages of the orders submitted by all traders in the market, and the right column reports the submission rate (SR), taking rate (TR), and execution rate (ER) for each type of traders and the market in whole. Based on Table 6, we obtain the following results on the effect of HFT and learning on the order choice and liquidity supply and consumption.[22]

First, informed HF traders submit less market orders but more aggressive limit orders and limit orders at the quote than LF traders, hence reduce liquidity consumption and increase liquidity supply, together with a significantly low execution. Overall, HFT improves both liquidity supply and consumption. This effect becomes more significant with learning and more competition among informed HF traders, however high volatility leads to an increase in both market and aggressive limit orders.

We now elaborate this result by comparing cases HF, NH, NL, 5H, 20H and LV. We first compare the order submission between HF and LF informed traders in cases HF and NH. In the HF case, compare to LF informed traders, HF traders submit less market orders (MO) and unaggressive limit orders (ULO), but more aggressive limit orders (ALO) and limit orders at the quote (LOA). Comparing to case NH, the informed HF traders submit even less market orders and more aggressive limit orders than LF traders. On the liquidity, both TR and SR for the informed HF traders are higher comparing to the LF traders in both HF and NH cases. This implies that the informed HF traders increase both liquidity supply and consumption. It is not surprised to see a low execution rate for the HFT, which implies that the HF traders submit and cancel their limit orders quickly. The increase in liquidity supply is due to the more aggressive limit orders from the informed HF traders, while the increase in liquidity consumption is mainly driven by high trading volume and high cancelation of the limit orders from the HFT, though the fraction of the market order from the informed HF traders is relative low to the aggressive limit orders.

[22]Note that, comparing cases HF to SL and IUH, we find that the effect of the information lag and uninformed HF traders is not significant.

Table 6. Order submission, liquidity supply and consumption, here the middle column represents, for each case, the percentages of four types of orders: MO, ALO, LOA and ULO for each type of traders: IH (informed HF), IL (informed LF), UL (uninformed LF) and for the whole market: All; while the right column represents the corresponding taking rate (TR), submission rate (SR), and the execution rate (ER).

Case	Trader	MO	ALO	LOA	ULO	TR	SR	ER
HF	IH	22.79	25.62	27.02	24.57	60.55	77.21	19.23
	IL	24.58	22.00	26.60	26.82	49.05	75.42	33.86
	UL	26.11	21.64	25.98	26.28	45.44	73.89	42.42
	All	24.77	23.16	26.40	25.66	50.00	75.23	32.68
NH	IL	25.83	16.68	28.61	28.88	51.05	74.17	33.39
	UL	28.14	15.91	27.89	28.06	49.90	71.86	39.32
	All	27.91	15.99	27.96	28.14	50.00	72.09	38.71
NL	IH	25.64	22.63	25.92	25.81	65.81	74.36	17.91
	IL	24.09	23.49	26.21	26.21	47.79	75.91	34.68
	UL	25.02	24.20	25.27	25.52	43.11	74.98	44.03
	All	25.20	23.57	25.57	25.67	50.00	74.80	33.69
5H	IH	26.25	24.12	26.75	22.89	62.83	73.75	21.06
	IL	24.59	21.44	26.81	27.16	50.12	75.41	32.46
	UL	26.73	20.58	26.13	26.56	46.85	73.27	41.39
	All	26.46	21.46	26.32	25.75	50.00	73.54	35.99
20H	IH	20.31	26.49	27.19	26.02	55.88	79.69	20.13
	IL	24.70	22.27	26.39	26.64	48.59	75.30	34.70
	UL	25.80	22.25	25.73	26.22	45.12	74.20	42.30
	All	22.74	24.58	26.56	26.12	50.00	77.26	29.43

On the learning effect, we compare HF and NL cases. With learning, the informed HF traders submit less market orders (MO by 2.85%=22.79%-25.64%) and unaggressive limit order (ULO by 1.24%=24.57%-25.81%), but more aggressive limit orders

Table 6 cont.

Case	Trader	MO	ALO	LOA	ULO	TR	SR	ER
3λ	IH	15.63	29.39	27.96	27.02	57.56	84.37	13.66
	IL	24.45	22.62	26.40	26.53	49.34	75.55	33.22
	UL	25.57	22.74	25.65	26.04	45.21	74.43	41.63
	All	20.08	26.38	26.95	26.60	50.00	79.92	25.13
12λ	IH	26.92	21.93	26.82	24.33	58.36	73.08	26.29
	IL	24.69	20.83	27.10	27.38	50.44	75.31	32.21
	UL	27.15	19.54	26.47	26.84	47.69	72.85	40.89
	All	26.93	20.18	26.60	26.29	49.99	73.07	36.87
48λ	IH	25.03	18.68	27.45	28.84	43.41	74.97	43.53
	IL	24.66	18.99	27.77	28.57	50.95	75.34	31.51
	UL	28.27	16.87	27.19	27.67	50.48	71.73	38.67
	All	27.74	17.18	27.26	27.83	49.99	72.26	38.39
LV	IH	21.76	22.35	29.17	26.72	56.26	78.24	21.62
	IL	25.47	19.90	27.09	27.56	48.72	74.54	35.97
	UL	27.09	18.74	26.83	27.33	47.29	72.91	41.42
	All	24.98	20.17	27.73	27.11	50.00	75.02	33.31
SL	IH	22.73	25.26	27.04	24.97	60.72	77.27	19.03
	IL	24.77	22.05	26.54	26.64	49.30	75.23	33.86
	UL	25.94	21.47	25.96	26.63	45.35	74.06	42.21
	All	24.85	22.05	26.49	26.60	50.00	75.34	32.75
IUH	IH	25.64	25.16	26.24	22.95	58.16	74.36	24.81
	UH	26.63	19.39	27.10	26.89	54.62	73.38	30.15
	IL	24.62	21.54	26.75	27.09	49.53	75.38	33.27
	UL	26.51	21.03	26.01	26.45	46.60	73.49	41.34
	All	26.26	21.53	26.30	25.91	50.00	73.74	35.62

(ALO by 2.99% =25.62%-22.63%) and limit orders at quote (by 1.10%=27.02%-25.92%). This is consistent with Chiarella et al. (2015), but becoming more significant with HFT, that the learning makes the informed HF traders use less market orders and more limit orders to gain better price advantage and to reduce the pick-off risk by cancelling their unexecuted limit orders quickly. It helps the informed HF traders to improve their profit opportunity, consistent with the result on order profit reported in the previous analysis. On the liquidity, with learning, TR reduces (from 65.81% to 60.55%) and SR increases (from 74.36% to 77.21%). The reduction in the taking rate and thus in liquidity consumption is due to the decrease in market order submission, while the increase in the submission rate and hence in the liquidity supply is driven by more aggressive limit orders submitted by the informed HF traders. However, comparing to LF traders, a higher TR for the informed HF traders indicates that HFT also improves liquidity consumption.

On the competition among the informed HF traders, we compare cases 5H, HF and 20H. When the number of informed HF traders increases from 5 to 10 and then 20, their market orders (MO) reduce from 24.85% to 21.76% and then to 20.36%, while their aggressive limit orders (ALO) increase from 21.72% to 22.35% and then to 22.71%, and unaggressive limit orders (ULO) also increase from 24.39% to 26.72% and then to 28.02%. This implies an increase in liquidity supply and a decrease in liquidity consumption when the competition among the informed HF traders becomes intensive. On the liquidity, as the number of the informed HT traders increases, TR reduces from 62.83% to 60.55% and then to 55.88%, while SR increases from 73.75% to 77.21% then to 79.69% for the informed HF traders. This is consistent with the reduction in market orders and increase in aggressive limit orders for the informed HF traders. This implies that competition among the informed HF traders reduces liquidity consumption but increases liquidity supply. Similar to the effect of the learning, comparing to LF traders, a higher TR for the informed HF traders in all the three cases indicates that HFT also improves liquidity consumption.

The effect of information uncertainty is however different. Comparing HF and HV cases, both the market and aggressive limit orders for the informed HF traders increase while their passive orders decreases as the fundamental volatility increases. Intuitively, when volatility is higher, the informed HF traders face more adverse selection by the change of fundamental value. Hence they place orders more aggressively to reduce the pick-off risk. On the liquidity, with higher volatility, TR increases while SR decreases for the informed HF traders. Therefore a high volatility makes informed HF traders submit more aggressive orders (MO and ALO) but less passive orders (LOA and ULO) and hence increases liquidity consumption and supply. Also, comparing to the LF traders, both TR and SR are higher for the informed HF traders, implying increase in both liquidity supply and consumption.

Secondly, the trading speed has a nonlinear effect on the order submission and liquidity supply and consumption, in particular, displaying a seemingly U-shaped relation to the aggressive limit orders and submission rate, and hence in liquidity supply, but a significant invert U-shaped relation to market orders and taking rate, and hence in liquidity consumption. This result is based on comparison among cases $3\lambda, HF, 12\lambda$ and 48λ. On order submission, with high arriving rates for the informed HF traders, comparing case HF to case 3λ, the market orders reduce from 22.79% to 15.63% and the aggressive limit orders increase from 25.62% to 29.39%. Hence market orders decrease while the aggressive limit orders increase with the speed of the HFT. This effect is also significant when we compare case 12λ to case HF. However, with a lower arriving rate $\lambda_h = 1/48$, comparing case 48λ to case 12λ, the market orders increase from 25.03% to 26.92% and the aggressive limit orders also increase from 18.68% to 21.93%. In these cases, market orders increase while the aggressive limit orders decrease with the speed of the HFT. For the limit orders at quote and unaggressive limit orders, with the increase of trading frequency, they decrease first from case 48λ to case 12λ, and then increase from case 12λ to case 3λ. On liquidity, from case HF to case 3λ, TR reduces from 60.55% to 57.56% and SR increases from 77.21% to 84.37%; while from case 48λ to case 12λ, TR increases from 43.41% to 58.36% and SR reduces from 74.97% to 73.08%. This nonlinear relation indicated a trade-off in liquidity supply and consumption with respect to the speed of HFT, which has an important policy implication in market design and liquidity.

Finally, informed HF traders make LF traders, both informed and uninformed, submit more aggressive limit orders(ALO) and less market orders (MO) and passive orders (LOA and ULO) and hence reduces their liquidity consumption and increase liquidity supply and executed rate, in particular for uninformed traders. This result is based on the comparison between cases NH and HF. On order submission, comparing case NH to case HF, the AlO of LF informed traders increases (from 16.68% to 22.00%), while their MO reduces (from 25.83% to 24.58%), and LOA and ULO reduce (from 57.49% =28.61%+28.88% to 53.42% =26.60%+26.82%). In the meanwhile, the ALO of LF uninformed traders increases (from 15.91% to 21.64%), while their MO reduce (from 28.14% to 26.11%), and their LOA and ULO reduce (from 55.91% =27.89%+28.06% to 52.26% =25.98%+26.28%). On the liquidity, comparing case NH to case HF, for uniformed LF traders, TR reduces (from 49.90% to 45.44%), while SR increases (from 71.86% to 73.89%), and ER also increases (from 39.32% to 42.42%). For informed traders, the effect is the same.

In summary, HFT affects the order submission behaviour and market liquidity significantly. In general HF traders tend to have a significantly low execution, submit less market orders but more aggressive limit orders and hence reduce liquidity consumption and increase liquidity supply. This effect becomes even more significant with learning and more competition among informed HF traders. However high volatility in the fundamental value leads to an increase in both market and aggressive limit orders and therefore to an increase in both liquidity supply and consumption. Also, the trading speed of HFT generates a seemingly U-shaped relation to aggressive limit orders and liquidity supply, but a significant invert U-shaped relation to market orders and liquidity consumption. Furthermore, informed HF traders make LF traders submit more aggressive limit orders(ALO) and less market orders (MO) and hence increase liquidity supply and reduce liquidity consumption. However, with high trading volume and low execution rate, HFT improves the overall market liquidity supply and consumption.

4.3. **Serial correlation and event clustering in order flows.** We now study the effect of the HFT on the serial correlation of order flows. Biais, Hillion and Spatt (1995) find that conditional order arrival frequencies have positive serial correlation. That is, the incoming order type is most likely to follow the same order type, such as a market buy order has higher probability of following a market buy order than other order type. Goettler, Parlour and Rajan (2005) also report significant positive serial correlations in order flows when traders have different private value to trade and they call this effect as "order persistence". If the conditional probabilities of all limit order types are significantly higher than corresponding unconditional probabilities, it is called *event clustering* effect documented in Gould, Porter, Williams, Fenn and Howison (2013). Based on the eight order types in Table 2, we report in Table 16 in the Appendix the conditional probabilities in percentage of the incoming order types following by the same order types, together with the corresponding unconditional probabilities. More importantly, we report the differences between the conditional probabilities and the unconditional probabilities to check the event clustering effect in Table 7.

TABLE 7. The event clustering, here D is the difference between CP and UCP.

Case	MB	ALB	LBA	ULB	MS	ALS	LSA	ULS
D_HF	3.13	5.15	6.87	6.37	3.46	5.78	6.79	6.45
D_NH	-0.24	-0.71	0.85	0.93	-0.21	-0.64	0.80	0.97
D_NL	2.46	2.25	4.27	4.39	2.52	2.34	4.26	4.37
D_5H	2.59	3.65	4.56	4.09	2.92	4.44	4.67	4.18
D_20H	2.91	5.38	9.68	9.42	3.09	5.78	9.42	9.52
D_3λ	4.37	10.14	12.57	12.59	4.97	10.19	12.02	12.85
D_12λ	2.07	2.37	3.70	3.51	2.15	2.94	3.76	3.57
D_48λ	-0.05	0.12	1.26	1.61	-0.03	0.41	1.36	1.57
D_LV	2.20	3.08	7.27	6.72	1.81	4.03	7.57	7.03
D_SL	3.28	5.04	6.85	6.4	3.52	5.86	6.79	6.48
D_IUH	2.29	2.79	3.75	3.73	2.72	3.34	3.80	3.75

Table 7 shows that in the HF case, the conditional probabilities of all order types are significantly higher than corresponding unconditional probabilities (say D_HF, all are significantly higher than 3%, for limit orders, the differences are higher than 5%). This implies that HFT increases the positive serial correlation of all the order types and generates event clustering effect. However, this is not always the case for the NH case (see the row for D_NH), some are negative and all are less than 1%. The positive serial correlation and the event clustering effect become more significant when trading frequency is higher. However, when the trading frequency is much lower, see case 48λ, the effect disappears. While without learning (see D_NL), lower volatility (see D_LV) and shorter information lag (see D_SL), the effect still holds. Therefore high speed in HFT contributes significantly to this effect.

4.4. **Information efficiency, market volatility and liquidity.** We finally study the impact of the HFT on the market, including the market information efficiency, volatility, and liquidity. We use the distance of the market price to the fundamental value to measure the market informational efficiency. Following Theissen (2000), we use Mean Absolute Error (MAE) to measure the absolute error of the market price from the fundamental value and Mean Relative Error (MRE) to measure the relative error of the market price from the fundamental value over the LF time periods,

$$MAE = \frac{1}{Y}\sum_{T=1}^{Y}|p_T - v_T|, \qquad MRE = \frac{1}{Y}\sum_{T=1}^{Y}\frac{|p_T - v_T|}{v_T}. \tag{1}$$

We also use the Kurtosis to measure fat tails in the market-price return, as well as the market efficiency, and the standard deviation (STD) of returns to measure the market volatility.

On the information efficiency and market volatility, we report the results in Table 8, showing that HFT improves information dissemination efficiency and hence price discovery, but increases the market volatility. Comparing case NH to case HF, MAE reduces from 36.82% to 7.39% and MRE reduces from 2.21% to 0.30%, which indicate that HFT significantly improves information dissemination efficiency. One reason is that a substantially increase in the competition among all the HD informed traders leads to releasing more information to the market. Moreover, comparing case HF to case NL, the MAE and MRE are larger, meaning that learning of the informed HF traders reduces information dissemination efficiency, which is consistent with Chiarella et al. (2015) that the learning of informed traders may help them to manipulate the order book and hence reduce the information efficiency. On the volatility, comparing case HF to case NH, the HFT increases the market volatility

Table 8. Market information efficiency and volatility, here MAE is in ticks and STD is in basis points.

Case	MAE	MRE	STD_{p_T}	$STD_{p_T^m}$
HF	7.39	0.30%	53.19	15.41
NH	36.82	2.21%	39.28	8.79
NL	4.94	0.19%	43.88	10.42
5H	9.57	0.39%	49.44	12.05
20H	8.41	0.31%	71.88	20.82
3λ	12.65	0.37%	85.86	19.06
12λ	8.86	0.38%	44.67	12.12
48λ	18.84	0.74%	35.09	8.41
LV	2.93	0.13%	27.31	10.62
SL	7.44	0.31%	55.14	15.49
IUH	11.47	0.39%	51.79	11.41

STD_{p_T} (in bps) by about 35% (from 39.28 to 53.19 bps) and $STD_{p_T^m}$ by about 75% (from 8.79 to 15.41 bps). This is consistent with Martinez and Rosu (2011) who study the impact of informed HFT, but different from Hagströmer and Nordén (2013) who examine the impact of market-making HFT, a different HFT setting from this paper.

In addition, comparing among other cases, we find that market efficiency reduces and volatility increases with increasing in the volatility of the fundamental value and decreasing in the number of the informed HF traders. Looking at case IUH, HF uninformed traders reduces the market efficiency as well as the volatility. The result for case SL indicates that the effect of the information lag is not significant. Interestingly, the speed of HFT has a nonlinear U-shaped relation to the information efficiency, but the market volatility always increases with the speed of HFT. This implies a trade-off between the HFT speed and market efficiency, which underlies the puzzles and debates of HFT on market efficiency.

TABLE 9. Market liquidity in terms of trading volume, the depth (D5) of the best five quotes on the sell side, and the bid/ask spread (in ticks).

Case	Volume	D5	Spread
HF	6.6	7.7	11.9
NH	4.7	13.2	6.2
NL	6.9	8.4	10.1
5H	5.7	8.4	10.7
20H	8.3	7.3	16.3
3λ	7.3	6.7	24.1
12λ	5.8	9.0	9.4
48λ	4.9	12.0	7.5
LV	6.6	8.6	5.2
SL	6.5	7.7	12.0
IUH	7.0	8.2	13.45

We also examine the impact of the HFT on the bid-ask spread, order book depth, and trading volume.[23] The results are reported in Table 9, showing that HFT increase the bid-ask spread and trading volume but reduces order book depth. Comparing to the NH case, HFT increases the trading volume by about 40% (from 4.7 to 6.6), which is consistent with Martinez and Rosu (2011). The order book depth of the best 5 quotes (on the sell side) is reduced by about 71% (from 13.2 to 7.7), which is consistent with Brogaard (2010) who find that HFT traders actually supply less market depth than other type traders. The bid-ask spread increases by about 92% (from 6.7 ticks to 11.9 ticks), which is consistent with Kim and Murphy (2013).

Furthermore, comparing cases 3λ, HF, 12λ and 48λ, we find that both volume and spread increase while the market depth decreases in the speed of HFT. This is consistent with the empirical study of Gai et al. (2012) who find that when the

[23]The trading volume is in unit for every minute. Hence total trading volume is equal to $4.7 \times 360(minutes) \times 100(days) = 179,200$ in NH case, which is approximately equal to the total MO submissions.

trading speed increases, HFT does not narrow the bid-ask spread and increase market depth. Comparing cases NL, LV, SL and IUH to HF, we find that learning increases the spread and reduces volume and depth. Comparing cases 5H, HF and 20H, we see that volume and spread increase while depth decreases with more informed HFT. Information uncertainty increases the spread and reduces the depth, but has no much impact on the trading volume. While the information lag has no significant impact, the uninformed HF traders increase volume, depth and spread. The active HFT is different from that under the passive market-making HFT, in which market makers only submit limit orders, which narrows the bid-ask spread and increases market depth under the HFT. Empirically, the impact of the HFT on the liquidity seems not clear, depending on whether the market making HFT dominates the market.

5. Conclusion

HFT is becoming a dominate trading in financial markets, but its impact on the markets is less clear. This paper provides a unified framework of market microstructure and learning literatures to examine the impact of learning and HFT in limit order markets. By employing the GA with a classifier system on market conditions, we allow traders to learn from market information including historical prices, fundamental value, quotes, the bid-ask spread, the order book imbalance and the last transaction sign. We also allow both high and low frequency traders in the market to examine their interaction and impact to each other. All the traders learn from the market, interact via the limit order book, and submit orders based on the market conditions.

We show that, with the GA learning, the informed HF traders use more market information, in particular the information related to the quotes and bid/ask spreads. The results show that it is the speed but more importantly information advantage and learning that generates profit opportunity for HFT. Compare to the LF informed traders, the informed HF traders submit less market orders but more aggressive limit orders, and increase both liquidity consumption and supply. In particular, the learning plays a more important role for informed HF traders' order submission, it makes informed HF traders reduce market orders, increase aggressive limit orders and unaggressive limit orders. Consequently, the HFT also leads to positive serial correlation and event clustering effect in order flows. On the impact to the market, we find that the informed HFT improves market information efficiency, increases market volatility, the bid-ask spread and trading volume, and reduces order book depth. We also examine the effect of competition among informed HFT, volatility of fundamental value, the information lag for uninformed traders, and the speed of HFT. In particular, we find that the speed of HFT is positively related to trading volume and spread but negatively related to market depth. More interestingly, the speed of HFT has an U-shaped relation to aggressive limit order submission, liquidity supply and market efficiency, but an invert U-shaped relation to HFT profit, market order submissions, and liquidity consumption. The results provide some insight on the trading behaviour and profitability of the HFT, its impact on market liquidity and order flow, and some implications on market policy and design..

The model proposed in this paper can be developed further in several directions. It is clear that learning plays more important role in algorithm and HF trading. This paper uses the GA with a classifier system as a learning mechanism, to which the literature on machine learning can contribute significantly (see, for example, Kearns and Nevmyvaka (2013)). To develop more important features or market information indicators and incorporate them into the classifier system for learning are important for the further development. Also, as point out in Easley et al. (2013), it is the "volume-clock" instead of "chronological time" that is essentially the core of HFT. Therefore, to develop HFT model based on the volume-clock is more desirable.

Appendix A: The design of trading rules and the GA Learning

In the Appendix A, we report the design of trading rules and the GA learning in Chiarella et al. (2015).

The design of trading rules in a GA with a classifier system. The learning mechanism of GA is based on the principles of natural selection. The outcome or solution of GA learning is called a chromosome, which is evaluated based on its historical performance and selectively evolved through processes of selection, crossover and mutation (to be defined in the latter half of this subsection). In the framework of Arthur et al. (1997), a classifier system is introduced so that an agent can recognize market conditions and choose the chromosome accordingly. For our model, a chromosome corresponds to a trading rule. Motivated by Arthur et al. (1997), we develop a classifier system to characterize market conditions and limit order book. A trading rule i contains two components. The first component corresponds to market conditions x^i (say, for example, the current mid-price $p^m_{t'}$ is higher than the expected fundamental value v^j_t of trader j) and the second component is an action y^i of buying or selling and order aggressiveness.[24] A trader then chooses the best trading rule according to its strength mainly determined by its historical performance (to be specified later). We now provide some details about the two components of a trading rule based on GA learning.

As the first component, the market condition x^i is based on classifier rules (CRs) of the classifier system, which is motivated by Goettler et al. (2009), Menkhoff, Osler and Schmeling (2010) and Wei et al. (2014). Goettler et al. (2009) find that the change in ask/bid, the last transaction price, the last transaction sign (buy or sell), the depths at the quote and away from the quote significantly affect the expectation of the fundamental value for uninformed traders. Menkhoff et al. (2010) find that the order submission of informed traders is affected by the bid-ask spreads, volatility, momentum of order flow and order book depth. The classifier system developed in the following extends the one introduced in Wei et al. (2014) who find that the

[24]The order aggressiveness is determined by the order type and order price, a market order is more aggressive than a limit order, and a limit order at quote is more aggressive than a limit order away from the quote.

forecasting accuracy of uninformed traders improves when they use the GA to learn from the lagged fundamental value, historical prices and the mid-price. In this paper, we use expected fundamental value, mid-price, historical prices and the order book information including recent change in quotes and the bid-ask spread, and order book depth imbalance to introduce 10 CRs listed in Table 1 to describe market conditions.

The CRs in Table 1 are grouped based on five aspects of market information. The first group "Fundamental Value" (FV) is related to fundamental value and contains the classifier rule CR1, which describes the relations between the expected fundamental value and the mid-price; the second group "Technical Rules" TR is related to technical rules, contains CR2 to CR4, and describes the technical rules among the mid-price, the expected fundamental value and the average market prices of the last $\tau/2$ and τ periods; the third group "Quotes and Spread" (QS) is related to the quotes and the spread, contains CR5 to CR7, and describes the recent change in the quotes and the bid-ask spread; the fourth group "Depth Imbalance" (DI) is related to the depth imbalance on the order book, contains CR8 to CR9, and describes the limit order book imbalance; and the last group Transaction Sign" TS is related to transaction sign of last market order, contains CR10, and describes the last transaction sign.[25] We use binary strings to represent CRs and hence market condition. For example, "1" indicates that CR1 is true and "0" means that CR1 is false. Hence one binary string has 10 bits and every bit represents two states of each CR, for example, "101110 01101" indicates one possible market condition. Hence, there are totally $2^{10} = 1024$ market conditions.

In principle, we have 1024 trading rules to match all the market conditions. However, in some cases, some market information become irrelevant for traders, and in such case we use "#" to replace 1 or 0, indicating that the corresponding market information is not considered. Trading rules with n "#" can match 2^n market

[25]If we let the agent consider all the information of the limit order book and prices of the past τ periods, the agent may learn better, but it leaves the set of classifier rules too large and the learning more complicated. In our model, agents are bounded rational, and they can process part of information of the limit order book. In the classifier system, the average prices and the order book depth reflect part of information of past τ periods.

conditions. For example, "##1#1 00#1#" represents a trading rule that matches $2^5 = 32$ market conditions. So a classifier rule and its corresponding market information may not be used in every trade. More realistically, traders have the limit ability to process all market information so their trading rules set is smaller than the set under full market conditions. To make the learning more efficient, we set the number of trading rules to be $2^8 = 256$. This means that some trading rules contain no less than two "#" in the market condition parts.

The second component of a trading rule is the action corresponding to buy/sell and order aggressiveness. In general, a trader needs to choose from many types of orders. Goettler et al. (2005) classify orders into four types, including market order, aggressive limit order, limit order at the quote, and limit order away from the quote. Similarly, in Menkhoff et al. (2010) orders are classifier into market orders, aggressive limit orders, and patient limit orders (limit orders at the quote and limit orders away from the quote). In this paper, we follow Goettler et al. (2005) and classify orders into four types: a market order (MO), a limit order at the quote (LOA), an aggressive limit order (ALO), and an unaggressive limit order (ULO, limit order away from the quote). To simplify the analysis, we define an aggressive limit order (ALO) to be the limit order above the bid or below the ask by one tick, and an unaggressive limit order (ULO) to be the limit order below the bid or above the ask by one tick.[26] Therefore an ALO narrows the bid-ask spread and improves the liquidity, while a LOA does not narrow the bid-ask spread but supplies immediate liquidity. Given the two sides of the book and the four types of orders, there are 8 actions in total, listed in Table 2. We use three binary bits to describe actions. For example "000" means a market buy (MB) order. For the informed traders, since they know the fundamental value, their buy/sell decision is determined by comparing the fundamental value to the bid and ask, and they can use GA to optimize their order aggressiveness. For the uninformed traders, they can use the GA to optimize both the buy-sell decision and order aggressiveness.

[26] Generally, the ALO or ULO can deviate from the quote in several ticks. To simplify the analysis, we set the deviation as one tick.

By combining the two components, a trading rule (x^i, y^i) means to take an action y^i under market conditions x^i. For example, one possible trading rule i can be defined when x^i is given by "1#1#1 00#1#" and y^i is given by "000". In some special limit order book scenarios, certain types of actions or orders are impossible or unused. For example, when the buy-side of the limit order book is empty, traders can not submit a market sell MS. These scenarios are listed in Table 10.

TABLE 10. The restrictions of actions.

Scenario		Unused action
The book is not empty	The bid-ask spread is more than one tick	None
	The bid-ask spread is equal to one tick	ALB and ALS
The book is empty	Only when the buy side is empty	MS, ALB & ULB
	Only when the sell side is empty	MB, ALS & ULS
	Both the buy and sell sides are empty	MB, ALB, ULB, MS, ALS & ULS

The GA learning. We consider the individual GA learning and the evolution process of the GA includes selection, crossover and mutation. In the selection process, a trading rule is selected by a tournament mechanism based on its strength $\eta^i = \pi^i - \delta^i$, where π^i represents the performance and δ^i measures the specificity of a trading rule.[27] The performance π^i of trading rule i for a trader when he enters the market at time t' is updated (with a zero initial value) according to

$$\pi^i_{t'} = \pi^i_z = \beta r^i_z + (1-\beta)\pi^i_{z-1}, \tag{2}$$

where $\beta \in [0, 1]$ and r^i_z is the order profit of rule i with $r^i_z = v_z - p_z$ for an executed buy order, $r^i_z = p_z - v_z$ for an executed sell order, or $r^i_z = 0$ for an canceled or expired order at the last trading time z. This means that the performance π^i of trading rule i for the trader is a weighted average of his recent order profit r^i_z and his previous performance π^i_{z-1} of the rule. A larger β means that traders weight more on the recent profit and less on the historical performance of the rule.

[27]The specificity measures the fitness or cost of a trading rule. For example, for a trading rule with "1#1#1 00#1#", the number of specific bits (non-omitted bits) m is equal to 5. The specificity of a trading rule is equal to $m\mu$, where μ is the bit cost, a small value such as 0.001. Hence, if two trading rules have the same performance π^i, the trading rule with less specificity (with more omitted bits, more adaptability) has a higher strength.

Initially, all the trading rules of a trader are randomly generated with most bits equalling to "#" in the market condition part.[28] With an initial performance of zero, the strength of a trading rule for the trader is low. When a submitted order has been executed, or canceled, or expired, the performance and hence the strength of the rule are updated by the trader. For the informed trader, the performances of his rules are updated immediately when he enters the market. For the uninformed traders, due to the information lag, the performances of his trading rules are updated only when the transactions occur before or at period $t - \tau$.

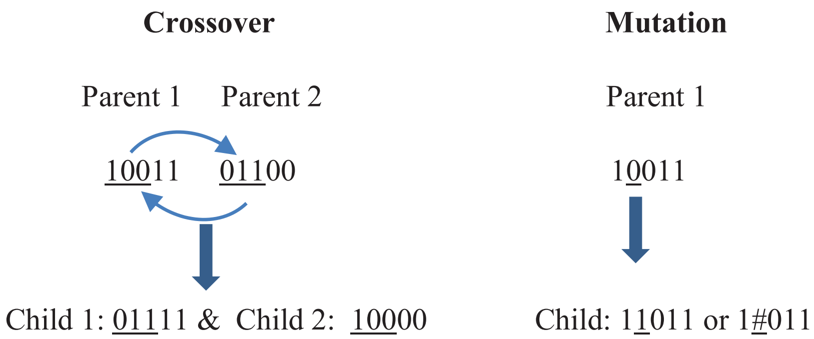


Figure 3. The crossover and mutation processes of a genetic algorithm.

When a trader enters the market, he ranks the performance of all his trading rules based on the strengths and selects ~~the top 10% of the rules~~ to generate new rules and replaces the bottom 10% of the rules. The new rules are generated through the processes of crossover and mutation according to given probabilities. Crossover means that, with a certain probability called the crossover rate, the trader ~~randomly~~ chooses two trading rules ~~from the top 10% of the rules~~ as parents, splits each trading rule into two parts at a random bit and then swaps the two parts to create two new trading rules as children. This process is illustrated in Figure 3. For example, the two parents trading rules are "10011" and "01100".[29] If they are split at the third

[28] Because the GA needs to use the historical data of the last τ periods (for generating technical rues), we let the market "warm up" for τ periods before traders use the GA to trade. In the warm up stage, traders randomly submit orders.

[29] In our GA with the classifier system, the condition part is a 10-bits string and the act part is a 3-bits string. To illustrate, we use 5-bits strings as an example.

bit, then the two new child rules are “01111” and “10000”. Mutation means that, with a certain probability called the mutation rate, the trader randomly selects a high strength trading rule as a parent and makes a random bit change of the parent trading rule to a different value. As illustrated in Figure 3, for the parent trading rule“10011”, the second bit is chosen to mutate, then the child rule becomes either “11011” or “1#011”. When the market condition parts crosses (mutates), the action parts also cross (mutate) with the same crossover (mutation) probability. The strength of the child trading rule is equal to the average strength of the parents under crossover and the strength of the parent under mutation (minus its specificity δ^i).

Appendix B: The usage frequency of classifier rule

Table 11. Usage frequency of CR groups of HFT traders.

CR group	Mean	Max	Min	STD	p
FV	0.4756	0.5003	0.4459	0.0088	0.49
TR	0.4731	0.4874	0.4559	0.0057	0.59
QS	0.4465	0.4673	0.4262	0.0078	0.20
DI	0.4849	0.5031	0.4649	0.0073	0.56
TS	0.4820	0.5017	0.4600	0.0078	0.46

Table 12. Usage frequency of informed traders.

CR	Mean	Max	Min	STD	p
FV	0.2740	0.2867	0.2588	0.0048	0.98
TR	0.3201	0.3316	0.3098	0.0037	0.11
QS	0.2879	0.2966	0.2797	0.0031	0.14
DI	0.3253	0.3348	0.3155	0.0039	0.36
TS	0.3361	0.3511	0.3221	0.0051	0.75

Table 13. Usage frequency of CR groups of uninformed traders.

CR	Mean	Max	Min	STD	p
FV	0.4093	0.4190	0.3991	0.0038	0.75
TR	0.4351	0.4413	0.4287	0.0025	0.51
QS	0.4903	0.4955	0.4851	0.0021	0.30
DI	0.4413	0.4479	0.4359	0.0022	0.36
TS	0.4247	0.4310	0.4165	0.0026	0.14

Appendix C: The numbers of order submission

Table 14 reports the total number of different types of orders for different types of traders, together with the aggregate numbers of order submissions in the market.

Table 14. Order submissions for each type of traders and aggregate order submissions. ELO is the executed limit orders, Total is the sum of all the submission orders.

Case	Trader	MO	ALO	LOA	ULO	ELO	Total
HF	IH	82,199	92,386	97,432	88,617	53,544	360,634
	IL	13,284	11,889	14,373	14,490	13,797	54,036
	UL	140,940	116,820	140,220	141,840	169,200	539,820
	All	236,423	221,095	252,025	244,947	236,541	954,490
NH	IL	15,500	10,010	17,170	17,330	14,860	60,010
	UL	152,010	85,950	150,660	151,560	152,640	540,180
	HFT	/	/	/	/	/	/
	All	167,510	95,960	167,830	168,890	167,500	600,190
NL	IH	92,430	81,601	93,451	93,070	48,014	360,552
	IL	13,014	12,690	14,157	14,157	14,220	54,018
	UL	135,090	130,680	136,440	137,790	178,290	540,000
	All	240,534	224,971	244,048	245,017	240,524	954,570
5H	IH	47,348	43,501	48,243	41,278	28,013	180,369
	IL	14,003	12,208	15,267	15,466	13,937	56,943
	UL	144,360	111,150	141,120	143,460	163,800	540,090
	All	205,711	166,859	204,629	200,204	205,749	777,402
20H	IH	146,560	191,112	196,168	187,720	115,726	721,560
	IL	11,880	10,712	12,696	12,816	12,568	48,104
	UL	139,320	120,150	138,960	141,570	169,470	540,000
	All	297,760	321,974	347,824	342,106	297,764	1,309,664

Table 15. Order submissions for each type of traders and aggregate order submissions(continue).

Case	Trader	MO	ALO	LOA	ULO	ELO	Total
3λ	IH	112,489	211,545	201,241	194,477	82,930	719,752
	IL	13,194	12,204	14,247	14,319	13,545	53,964
	UL	138,060	122,760	138,510	140,580	167,310	539,910
	All	263,743	346,509	353,998	349,376	263,785	1,313,626
12λ	IH	48,567	39,553	48,381	43,888	34,650	180,389
	IL	13,293	11,214	14,589	14,742	13,059	53,838
	UL	146,610	105,480	142,920	144,900	160,830	539,910
	All	208,470	156,247	205,890	203,530	208,539	774,137
48λ	IH	11,246	8,393	12,333	12,955	14,662	44,927
	IL	13,320	10,260	15,003	15,435	12,825	54,018
	UL	152,640	91,080	146,790	149,400	149,760	539,910
	All	177,206	109,733	174,126	177,790	177,247	638,855
$\cancel{2\kappa}$	IH	78,489	80,638	105,216	96,383	61,017	360,726
	IL	13,743	10,737	14,616	14,873	14,468	53,969
	UL	146,295	101,205	144,900	147,600	163,080	540,000
	All	238,527	192,580	264,732	258,855	238,564	954,694
$\cancel{180\tau}$	IH	81,994	91,115	97,538	90,069	53,033	360,716
	IL	13,365	11,898	14,319	14,373	13,743	53,955
	UL	140,040	115,920	140,130	143,730	168,750	539,820
	All	235,399	218,933	251,987	248,172	235,526	954,491
IUH	IH	46,296	45,431	47,374	41,441	33,305	180,541
	UH	48,042	34,981	48,903	48,512	39,911	180,437
	IL	13,266	11,610	14,418	14,598	13,518	53,892
	UL	143,100	113,490	140,400	142,740	163,980	539,730
	All	250,704	205,512	251,095	247,290	250,714	954,600

Appendix D: The conditional order submission

Table 16. Conditional order submission. Here CP is the conditional probability of the incoming order type following the same order type, and UCP is the unconditional probability.

Case	MB	ALB	LBA	ULB	MS	ALS	LSA	ULS
CP_HF	15.54	16.66	19.91	19.19	15.82	17.43	20.15	19.30
CP_NH	13.67	7.32	14.71	14.95	13.8	7.32	14.9	15.09
CP_NL	15.02	14.06	16.95	17.21	15.16	14.09	17.15	17.23
CP_5H	15.78	14.44	17.47	16.97	16.20	15.11	18.08	17.06
CP_20H	14.32	17.52	22.66	22.30	14.42	18.22	23.00	22.77
CP_3λ	14.43	23.17	25.78	25.69	14.99	23.54	25.75	26.34
CP_12λ	15.41	12.60	16.71	16.64	15.74	12.88	17.35	16.74
CP_48λ	13.74	8.84	14.59	15.43	13.92	8.87	15.30	15.57
CP_$\cancel{2\kappa}$	14.63	13.31	20.72	20.07	14.37	13.99	21.84	20.78
CP_$\cancel{180\tau}$	15.57	16.57	19.69	19.35	15.89	17.27	20.34	19.54
CP_IUH	15.35	13.89	16.59	16.56	15.83	14.26	17.08	16.60
UCP_HF	12.41	11.51	13.04	12.82	12.36	11.65	13.36	12.85
UCP_NH	13.91	8.03	13.86	14.02	14.01	7.96	14.10	14.12
UCP_NL	12.56	11.81	12.68	12.82	12.64	11.75	12.89	12.86
UCP_5H	13.19	10.79	12.91	12.88	13.28	10.67	13.41	12.88
UCP_20H	11.41	12.14	12.98	12.88	11.33	12.44	13.58	13.25
UCP_3λ	10.06	13.03	13.21	13.10	10.02	13.35	13.73	13.49
UCP_12λ	13.34	10.23	13.01	13.13	13.59	9.94	13.59	13.17
UCP_48λ	13.79	8.72	13.33	13.82	13.95	8.46	13.94	14.00
UCP_$\cancel{2\kappa}$	12.43	10.23	13.45	13.35	12.56	9.96	14.27	13.75
UCP_$\cancel{180\tau}$	12.29	11.53	12.84	12.95	12.37	11.41	13.55	13.06
UCP_IUH	13.06	11.10	12.84	12.83	13.11	10.92	13.28	12.85

References

Allen, T. W. and Carroll, C. D. (2001), 'Individual learning about consumption', *Macroeconomic Dynamics* **5**(2), 255–271.

Arifovic, J. (1994), 'Genetic algorithm learning and the cobweb model', *Journal of Economic Dynamics and Control* **18**, 3–28.

Arifovic, J. (1996), 'The behavior of the exchange rate in the genetic algorithm and experimental economies', *Journal of Political Economy* **104**, 510–541.

Arthur, W., Holland, H., LeBaron, B., Palmer, R. and Tayler, P. (1997), *Asset pricing under endogenous expectations in an artificial stock market*, in *The Economy As An Evolving Complex System II*, Eds. W. Arthur and S. Durlauf and D. Lane, Addison-Wesley, pp. 15–44.

Biais, B., Foucault, T. and Moinas, S. (2011), Equilibrium high frequency trading, Working paper, SSRN.

Biais, B., Hillion, P. and Spatt, C. (1995), 'An empirical analysis of the limit order book and the order flow in the paris bourse', *Journal of Finance* **50**, 1655–1689.

Brogaard, J. A. (2010), High frequency trading and its impact on market quality, Working paper, SSRN.

Brogaard, J. A. (2012), High frequency trading and volatility, Working paper, SSRN.

Brogaard, J., Hendershott, T., and Riordan, R. (2014), 'High-frequency trading and price discovery', *The Review of Financial Studies* **27**, 2267–2306.

Chiarella, C., He, X. and Wei, L. (2015), 'Learning, information processing and order submission in limit order markets', *Journal of Economic Dynamics and Control* **61**, 245–268.

Easley, D., de Prado, M. L. and OHara, M. (2013), *High Frequency Trading—New Realities for Traders, Markets and Rgulators*, Risk Books, chapter The Volume Clock: Insights into the High-Frequency Paradigm.

Gai, J., Yao, C. and Ye, M. (2012), The externalities of high-frequency trading, Working paper, SSRN.

Goettler, R. L., Parlour, C. A. and Rajan, U. (2005), 'Equilibrium in a dynamic limit order market', *Journal of Finance* **10**, 2149–2192.

Goettler, R. L., Parlour, C. A. and Rajan, U. (2009), 'Informed traders and limit order markets', *Journal of Financial Economics* **93(1)**, 67–87.

Gould, M., Porter, M., Williams, S., Fenn, D. and Howison, S. D. (2013), 'Limit order books', *Quantitative Finance* **13**(11), 1709–1742.

Hagströmer, B. and Nordén, L. (2013), 'The diversity of high-frequency traders', *Journal of Financial Markets* **16**, 741770.

Hasbrouck, J. (2013), High frequency quoting: Short-term volatility in bids and offers, Working paper, SSRN.

Hendershott, T., Jones, C. and Menkveld, A. (2011), 'Does algorithmic trading increase liquidity?', *The Journal of Finance* **66**, 133.

Hoffmann, P. (2014), 'A dynamic limit order market with fast and slow traders', *Journal of Financial Economics* **113**, 156169.

Holland, J. H. (1975), *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, University of Michigan Press, Ann Arbor, MI.

Kearns, M. and Nevmyvaka, Y. (2013), *High Frequency Trading—New Realities for Traders, Markets and Rgulators*, Risk Books, chapter Machine learning for market microstructure and high frequency trading.

Kim, S. and Murphy, D. (2013), The impact of high-frequency trading on stock market liquidity measures, Working paper, SSRN.

Kirilenko, A. A. and Lo, A. W. (2013), Moore's law vs. murphy's law: Algorithmic trading and its discontents, Working paper, SSRN.

Kirilenko, A., Samadi, M., Kyle, A. S. and Tuzun, T. (2011), The flash crash: The impact of high frequency trading on an electronic market, Working paper, SSRN.

LeBaron, B., Arthur, W. and Palmer, R. (1999), 'Time series properties of an artificial stock market.', *Journal of Economic Dynamics and Control* **23**, 1487–1516.

Lettau, M. and Uhlig, H. (1999), 'Rules of thumb versus dynamic programming', *American Economic Review* **89**, 148–174.

Marimon, R., McGrattan, E. and Sargent, T. J. (1990), 'Money as a medium of exchange in an economy with artificially intelligent agents', *Journal of Economic Dynamics and Control* **14**, 329–373.

Martinez, V. and Rosu, I. (2011), High frequency traders, news and volatility, Working paper, SSRN.

Menkhoff, L., Osler, C. L. and Schmeling, M. (2010), 'Limit-order submission strategies under asymmetric information', *Journal of Banking & Finance* **34**, 2665–2677.

Routledge, B. (1999), 'Adaptive learning in financial markets', *Review of Financial Studies* **12**, 1165–1202.

Routledge, B. (2001), 'Genetic algorithm learning to choose and use information', *Macroeconomic Dynamics* **5**, 303–325.

SEC (2010), Concept release on equity market structure, Release no. 34-61358; fileno. s7-02-10, U.S. Securities and Exchange Commission.

Theissen, E. (2000), 'Market structure, informational efficiency and liquidity: An experimental comparison of auction and dealer markets', *Journal of Financial Markets* **3(4)**, 333–363.

Vuorenmaa, T. A. and Wang, L. (2014), An agent-based model of the flash crash of may 6, 2010, with policy implications, Working paper, SSRN.

Wei, L., Zhang, W., He, X. and Zhang, Y. (2014), The value of learning and liquidity in limit order markets, Working paper(no. 333), Quantitative Finance Research Central (QFRC), University of Technolgy, Sydney.

Yan, Y. and Zhang, S. (2012), 'An improved estimation method and empirical properties of the probability of informed trading', *Journal of Banking & Finance* **36**(2), 454–467.

Zhang, X. (2010), High frequency trading, stock volatility, and price discovery, Working paper, SSRN.