

HEC MONTRÉAL
École affiliée à l'Université de Montréal

Three essays on high-frequency trading algorithms

par
Gabriel Yergeau

Thèse présentée en vue de l'obtention du grade de Ph. D. en administration
(option Finance)

Décembre 2017

© Gabriel Yergeau, 2017

HEC MONTRÉAL

École affiliée à l'Université de Montréal

Cette thèse intitulée :

Three essays on high-frequency trading algorithms

Présentée par :

Gabriel Yergeau

a été évaluée par un jury composé des personnes suivantes :

Tolga Cenesizoglu

HEC Montréal

Président-rapporteur

Georges Dionne

HEC Montréal

Directeur de recherche

Jean-Guy Simonato

HEC Montréal

Membre du jury

Ryan Riordan

Queen's University

Examineur externe

Nicolas Vincent

HEC Montréal

Représentant du directeur de HEC Montréal

Résumé

Cette dissertation cumulative étudie empiriquement les comportements d'algorithmes de négociation actifs sur les marchés boursiers. Nous présentons trois papiers examinant différentes questions de recherche. Malgré la diversité des contextes, un thème commun s'articule autour de l'évaluation de spécificités de la microstructure des marchés sur la présence et le comportement des négociateurs algorithmiques.

Le premier chapitre propose un algorithme de négociation à haute fréquence agissant à titre de mainteneur de marché endogène. Les décisions de fournir de la liquidité sont basées sur un modèle de gestion dynamique d'inventaire. L'algorithme inclut les caractéristiques de gestion de risque. Nous proposons une méthodologie de priorisation des ordres adaptée aux spécificités des données publiques de la bourse de Francfort. L'évaluation de la performance tient compte du dynamisme du levier financier et de la nature asynchrone de l'environnement à haute fréquence. Le comportement de l'algorithme lors des événements extrêmes ultra-rapides est analysé.

Le deuxième chapitre s'intéresse à l'identification d'empreintes algorithmiques à partir d'information publique. Notre approche s'articule autour de deux axes. Premièrement, nous surveillons les jaillissements d'activités à haute fréquence à l'aide d'une structure de données adaptée à un environnement de faible latence. Nous tenons compte de la distortion temporelle dynamique. Nous documentons la présence de bruit crypté. Deuxièmement, nous identifions la présence de négociateurs informés à partir de caractéristiques comportementales. Ceux-ci utilisent des algorithmes à haute et à basse latence. Nous associons 25% de tous les événements des enchères par lots à 9 types d'algorithmes.

Dans le troisième chapitre, nous utilisons l'algorithme conceptualisé au Chapitre 1 afin de comparer la performance d'un mainteneur de marché endogène sous deux scénarios soit la norme internationale de la priorisation temps-prix et la préférence du courtier, une spécificité de la microstructure des marchés canadiens. Nous considérons le modèle de

marché appliqué par le Toronto Stock Exchange ainsi que les imperfections de marché. Nous identifions les variables systématiques et idiosyncratiques influençant les positions de l'algorithme. Nous analysons le risque de contagion d'illiquidité, une préoccupation des investisseurs institutionnels et des législateurs.

Nous identifions la présence de certains algorithmes lors des séances d'enchères. Ceux-ci ont des comportements hétérogènes qui peuvent avoir des impacts diamétralement opposés sur le processus de découverte de prix. La conceptualisation et l'implémentation en sol allemand et canadien d'un algorithme destiné à un mainteneur de marché endogène permet d'établir la viabilité économique de ce type d'opérations. Malgré l'absence d'obligation formelle imposée par les Bourses de fournir de la liquidité, l'examen des comportements lors de périodes de forte volatilité ne nous a pas permis de documenter le retrait de l'offre de liquidité.

Mots clés : carnet d'ordres limites, enchère, finance comportementale, négociation algorithmique, négociation à haute fréquence, mainteneur de marché, microstructure.

Abstract

This cumulative dissertation empirically studies the behavior of trading algorithms active on stock markets. We present three papers examining different research questions. Despite the diversity of contexts, a common theme revolves around the assessment of the markets microstructure specificities on the presence and behavior of algorithmic traders.

The first chapter proposes a high-frequency trading algorithm acting as an endogenous market maker. The decisions to provide liquidity are based on a dynamic inventory management model. The algorithm includes risk management features. We propose a methodology of order prioritization adapted to the specificities of the Frankfurt stock exchange's public data. The performance assessment considers the leverage dynamism and the asynchronous nature of the high-frequency environment. The behavior of the algorithm during ultra-fast extreme events is analyzed.

The second chapter deals with the identification of algorithmic imprints from public data. We base our approach on two axes. First, we monitor high-frequency activity bursts using a data structure adapted to a low-latency environment and we consider dynamic time warping. We document the presence of encrypted noise. Second, we identify the presence of informed traders based on behavioral characteristics. These use high and low latency algorithms. Overall, we link 25% of all batch auctions events to 9 algorithm types.

In the third chapter, we use the algorithm conceptualized in Chapter 1 to compare the performance of an endogenous market maker under two scenarios: the international standard of time-price prioritization and broker preference, a microstructure specificity of Canadian markets. We consider the market model applied by the Toronto Stock Exchange as well as market imperfections. We identify the idiosyncratic and systematic variables influencing the algorithm positions. We analyze the risk of illiquidity contagion, a preoccupation of institutional investors and legislators.

We identify the presence of some algorithms during the auction sessions. These have heterogeneous behaviors that can have diametrically opposite impacts on the price

discovery process. The conceptualization and implementation in Germany and Canada of an algorithm intended for an endogenous market maker makes it possible to establish the economic viability of this type of operations. Despite the absence of a formal obligation on the stock exchanges to provide liquidity, examination of the behavior during periods of high volatility did not allow us to document the withdrawal of the liquidity offer.

Keywords : algorithmic trading, auction, behavioral finance, high-frequency trading, limit order book, market-making, microstructure.

Table of contents

Résumé	v
Abstract	vii
Table of contents	ix
List of tables	xi
List of figures	xiii
Abbreviations	xiv
Remerciements	xvii
Chapter 1 Profitability and Market Quality of High Frequency Market-makers: An Empirical Investigation	1
Abstract	1
1.1 Introduction	1
1.2 Optimal Quoting Policy of Liquidity Provision	3
1.2.1 OQP determination	5
1.2.2 Emulation and OQP	6
1.3 Empirical investigation	6
1.4 Data	9
1.5 TVWAR: a time- and volume-weighted average return	12
1.6 Results	15
1.6.1 Ait-Sahalia and Saglam (2014) model	16
1.6.2 Trading strategy	17
1.7 Robustness tests	20
1.7.1 Speed's impact on performance	20
1.7.2 Strategy's features	21
1.8 Conclusion	22
References	24
Appendix 1.1: Threshold calculation: an efficient algorithm	27
Chapter 2 Identification of algorithmic imprints using public data: an application to batch auctions	42
Abstract	42
Introduction	42
2.1 Literature review	44
2.1.1 Behavior	44
2.1.2 Machine learning	45
2.2 Institutional context: trading on Xetra	46
2.3 Algorithmic sequences	47
2.3.1 Negative price loops (NPLs)	47
2.3.2 Behavior-based strategies	49
2.4 Data	50
2.4.1 Transaction costs	51
2.5 Results	52
2.5.1 Negative price loops	52
2.5.2 NPLs and liquidity	54
2.5.3 Behavioral-based strategies	56
2.6 Discussion	56

2.7	Conclusion	59
	References	60
Chapter 3 Performance and behavior of endogenous liquidity providers		81
	Abstract	81
	Introduction	81
3.1	Trading on the TSX.....	84
3.2	Data.....	85
3.3	Endogenous liquidity provider	86
3.3.1	Optimal quoting policy and circuit-breakers	86
3.3.2	Closing of positions and market imperfections.....	88
3.3.3	Scenarios.....	88
3.4	Simulation of activities	89
3.5	Results.....	89
3.5.1	Sustainability of performances	90
3.5.2	Stylized facts.....	92
3.5.3	Influential market and stocks characteristics	93
3.5.4	Illiquidity contagion risk.....	95
3.6.	Conclusion	97
	References	99
Appendix	3.1 Optimal quoting policy.....	104
Appendix	3.2 Observed parameters value.....	107

List of tables

TABLE 1.1 MARKET SUMMARY STATISTICS, TRADES AND LOB1	28
TABLE 1.2 TWO-WAY CLASSIFICATION OF PRICE MOVEMENTS IN CONSECUTIVE INTRADAY TRADES: SUMMARY (,000)	28
TABLE 1.3 TWO-WAY CLASSIFICATION OF PRICE MOVEMENTS IN CONSECUTIVE INTRADAY TRADES DAX (%).....	29
TABLE 1.4 ULTRAFast EXTREME EVENTS (UEES) SUMMARY.....	30
TABLE 1.5 SIGNAL AND TRADE INDEPENDENCE: CHI-SQUARE TESTS	31
TABLE 1.6 DAILY AND INTRADAY PROFITABILITY – OQP	32
TABLE 1.7 ORDERS AND POSITIONS: AN EXAMPLE	33
TABLE 1.8 DAILY AND INTRADAY PROFITABILITY – TRADING STRATEGY	34
TABLE 1.9 HFMM’S TRADE ORIGINS.....	35
TABLE 1.10 IMPACT OF LEVERAGE ON PERFORMANCE	36
TABLE 1.11 PARTICIPATION IN TRADES	37
TABLE 1.12 IMPACT OF LATENCY ON PERFORMANCE	37
TABLE 1.13 IMPACT FROM STRATEGY'S FEATURES.....	38
 TABLE 2. 1 CLASSIFICATION OF TRADERS.....	65
TABLE 2. 2 XETRA MARKET MODEL.....	65
TABLE 2. 3 AUCTIONS: PUBLIC INFORMATION.....	66
TABLE 2. 4 INTERPRETATION OF AUCTION'S PUBLIC INFORMATION	66
TABLE 2. 5 BEHAVIOR-BASED TYPES OF ALGORITHM	67
TABLE 2. 6 STATISTICS - ALL AUCTIONS.....	67
TABLE 2. 7 INDICATIVE QUANTITY - ALL AUCTIONS	68
TABLE 2. 8 NPLS ALGORITHM TYPES	68
TABLE 2. 9 NEGATIVE PRICE LOOPS AND SWT LEVELS	69
TABLE 2. 10 NPLS ALGORITHMIC SEQUENCES	69
TABLE 2. 11 STABLE FAMILY PARAMETERS FEBRUARY – JULY 2013	70
TABLE 2. 12 KOLMOGOROV-SMIRNOV, TWO-SAMPLE TESTS.....	71
TABLE 2. 13 BEHAVIOR-BASED ALGORITHMIC EVENTS.....	71
TABLE 2. 14 BEHAVIOR-BASED TYPES OF ALGORITHM	72
TABLE 2. 15 DAX BEHAVIOR-BASED SEQUENCES : STATISTICS BY AUCTION FEBRUARY – JULY 2013.....	72

TABLE 2. 16 MDAX BEHAVIOR-BASED SEQUENCES : STATISTICS BY AUCTION FEBRUARY – JULY 2013.....	73
TABLE 3. 1 SESSIONS ON THE TSX.....	113
TABLE 3. 2 TSX: MAKER-TAKER REBATE AND FEE LEVELS.....	113
TABLE 3. 3 TRADES AND QUOTES : MARCH - AUGUST 2015	114
TABLE 3. 4 GLOBAL PERFORMANCE, PRICE-TIME PRIORITY VS BROKER PREFERENCE MARCH-AUGUST 2015	114
TABLE 3. 5 MONTHLY PERFORMANCE : PRICE-TIME AND BROKERS A AND B.....	115
TABLE 3. 6 STABLE DISTRIBUTION PARAMETERS : BROKERS A AND B PROFITS & LOSSES MARCH – AUGUST 2015.....	116
TABLE 3. 7 MARKET MAKER INDEX: BROKERS A AND B.....	116
TABLE 3. 8 QUOTES AND TRADES : BROKERS A AND B.....	117
TABLE 3. 9 REGRESSIONS ON FACTORS DETERMINING ELP'S VARIATIONS OF POSITION	118
TABLE 3. 10 BROKER B LEVEL 1 BID (UPPER PART OF) AND OFFER (LOWER PART OF) CROSS CORRELATION MATRIX DURING THE EXTREME EVENTS.....	119
TABLE 3. 11 BROKER B LEVEL 1 BID (UPPER PART OF) AND OFFER (LOWER PART OF) CROSS CORRELATION MATRIX DURING THE EXTREME EVENTS.....	120
A3.2. 1 TRADES BY MINUTE FIRST 5 DAYS, MARCH 2015.....	107
A3.2. 2 SIGNALS BY MINUTE, FIRST 5 DAYS, MARCH 2015.....	109
A3.2. 3 SPREADS BY MINUTE FIRST 5 DAYS, MARCH 2015.....	111

List of figures

FIGURE 1. 1 DAX DAILY QUOTES - FEBRUARY TO JULY 2013.....	39
FIGURE 1. 2 NUMBER OF UEEs PER DAY: DAX - MDAX	39
FIGURE 1. 3 NUMBER OF UEEs PER MINUTE: DAX - MDAX	40
FIGURE 1. 4 DAX - MDAX CUMULATIVE P&L: OQP.....	40
FIGURE 1. 5 DAX- MDAX LEVERAGED RETURN PER TIME INTERVAL.....	41
FIGURE 2. 1 SHIFTED WAVELET TREE - STRUCTURE	74
FIGURE 2. 2 DAX COMPONENTS - CLOSING AUCTIONS	75
FIGURE 2. 3 DEUTSCHE BANK 20130212 CLOSING AUCTION.....	76
FIGURE 2. 4 IDENTIFICATION PARAMETERS FOR NEGATIVE PRICE LOOPS	77
FIGURE 2. 5 NPL EVENTS BY 5-SECOND INTERVAL - CLOSING AUCTIONS : AGGREGATED STATISTICS FEBRUARY – JULY 2013	78
FIGURE 2. 6 NPL EVENTS BY 10-SECOND INTERVAL – CLOSING AUCTIONS : TYPE OF ALGORITHMS FEBRUARY – JULY 2013	79
FIGURE 2. 7 TRANSACTION COST OF NPL SEQUENCES	80
FIGURE 2. 8 TWAP EXAMPLE - VOLKSWAGEN.....	80
FIGURE 3. 1 VIX S&P/TSX 60, XIU.TO : MARCH - AUGUST 2015.....	121
FIGURE 3. 2 CUMULATIVE PROFIT EXCLUDING REBATES: BROKER PREFERENCE WITH BROKERS A AND B VS. TIME-PRICE PRIORITY.....	121
FIGURE 3. 3 CRITICAL MARKET STATES: REPARTITION BY NUMBER OF STOCKS INVOLVED : MARCH – AUGUST 2015	122
FIGURE 3. 4 CRITICAL MARKET STATES: REPARTITION BY MONTHLY OCCURENCES : MARCH – AUGUST 2015	122
FIGURE 3. 5 AVERAGE LEVEL 1 OFFER QUANTITY	123
FIGURE 3. 6 AVERAGE LEVEL 1 BID QUANTITY	123
FIGURE 3. 7 CROSS-CORRELATION MATRIX DISTRIBUTION	124

Abbreviations

AT algorithmic trading
CB(s) circuit-breaker(s)
DB Deutsche Boerse AG
DMA direct market access
DMMs designated market maker(s)
DTW dynamic time warping
ELP(s) endogenous liquidity provider(s)
EOD end of the day
HFMMs high frequency market maker(s)
HFTer(s) high frequency trader(s)
HFT high frequency trading
ID unique identifier
IIROC Investment Industry Regulatory Organization of Canada
LO(s) limit order(s)
LOBs limit order book(s)
LFT(s) low frequency trader(s)
MDP Markov decision process
NPL(s) negative price loop(s)
 μ S microsecond
OQP(s) optimal quoting policy(ies)
PnL(s) profit and loss statement(s)
SWT shifted wavelet tree
TSX Toronto Stock Exchange
TVWAR time-volume weighted average return
UEE(s) ultrafast extreme event(s)
UTD(s) update(s)

À Jocelyne ma compagne de vie. Elle m'a toujours montré de l'amour, de la patience et du soutien. Les nombreux sacrifices qu'elle a fait m'ont permis de compléter ce périple.

Remerciements

Tout d'abord, j'aimerais remercier mon superviseur Georges Dionne pour avoir accepté de me guider dans ce processus. Sa disponibilité et son expérience de recherche se sont avérées extrêmement précieux au cours de toutes ces années. Je le remercie de sa patience pour les très nombreuses lectures de mes papiers. Ses commentaires et ses idées ont su me faire progresser. Je le remercie pour le généreux soutien à mon financement.

Je remercie également mes collègues Yann Bilodeau, Stéphane Galzin, Mathieu T.-Blais et Zhou Xiaozhou de l'équipe de haute-fréquence de la Chaire de recherche du Canada en gestion des risques dont le professeur Dionne est le titulaire. Les rencontres nous permettant d'échanger sur nos travaux respectifs et leurs nombreux commentaires m'ont été fort importants.

Chapter 1

Profitability and Market Quality of High Frequency Market-makers: An Empirical Investigation

Abstract

Financial markets in contemporary regulatory settings require the presence of high-frequency liquidity providers. We present an applied study of the profitability and the impact on market quality of an individual high-frequency trader acting as a market-maker. Using a sample of sixty stocks over a six-month period, we implement the Ait-Sahalia and Saglam (2014) optimal quoting policy (OQP) of liquidity provision from dynamic inventory management model. The OQP allows the high-frequency trader to extract a constant annuity from the market but its profitability is insufficient to cover the costs of market-making activities. The OQP is embedded in a trading strategy that relaxes the model's constraint on the quantity traded. Circuit-breakers are implemented and market imperfections are considered. Profits excluding maker-fees and considering transaction fees are economically significant. We propose a methodology to adjust the returns for asynchronous trading and varying leverage levels associated with dynamic inventory management. This allows us to qualify high trade volume as a proxy of informed trading. The high-frequency trader behaves as a constant liquidity provider and has a positive effect on market quality even in periods of market stress.

1.1 Introduction

Most stock exchanges have removed or have diluted the formal obligation to maintain an orderly market once imposed on human market-makers: high-frequency liquidity suppliers are major participants in electronic markets (Anand and Venkataraman (2013); Menkveld (2013)). Jones (2013) explains the increase in high-frequency market making by lower cost structures and more adequate responses to adverse selection.

Notwithstanding the importance of high-frequency market-making, very little is known about the profitability and individual behavior of high-frequency liquidity providers. Menkveld (2013) describes and evaluates the activities of a large high-frequency market maker (HFMM) who uses spatial arbitrage as the core of his market-making strategy. He asserts that fees are a substantial part of the HFMM's profit and loss account. Serbera and Paumard (2016) argue that maker-fees represent the core profitability of high-frequency market-making. Popper (2012) states that profits in American stocks from high-speed trading in 2012 are down 74 percent from the peak of \$4.9 billion in 2009. This can be linked to a decrease in commission and rebates, reported by Malinova and Park (2015).

To deal with this trend, we assess the economic viability, excluding maker fees, of a typical HFMM in two steps: first, we emulate the behavior of an HFMM using Ait-Sahalia and Saglam (2014) dynamic inventory management model (the model hereafter). Their model mimics the high frequency trading stylized facts. Their setup differs from the classical dynamic inventory models (Grossman and Stiglitz (1980); Roll (1984); Glosten and Milgrom (1985); Kyle (1985)) in that the strategic variable is whether or not to quote rather than change the supply curve. It yields to an optimal quoting policy (OQP) of liquidity provision that drives the HFMM's trading decisions. Second, we embed the OQP in a trading strategy that relaxes some of the model's assumptions and adds risk management features.

Market-making implies asynchronous trades and varying market risk related to the dynamic leverage from the model's OQP and the liquidity demanders' needs. These factors affect the HFMM's trading performance. We propose a measure, the time-volume weighted average return (TVWAR), to cope with both phenomena. It allows us to analyze a single stock performance incurring different phases of trading activities and/or liquidity depth, and to compare returns of stocks with different idiosyncratic characteristics.

The emulation provides insights into the implications of high-frequency market-making on market quality, a matter that has raised much concern. Duffie (2010) describes the importance of monitoring the pattern of response to supply and demand shocks for asset pricing dynamics. Foucault, Kadan et al. (2013) develop a model based on an endogenous

reaction time to trading activities, and find that algorithmic trading plays an important role in monitoring the state of liquidity cycles. Biais, Foucault et al. (2015) and Pagnotta and Philippon (2015) analyze competition on speed. They argue that competition should have a positive effect on the price discovery process. Finally, market stability is documented using Johnson, Zhao et al. (2013) ultrafast extreme events (UEEs).

The paper is organized as follows. Section 1.2 presents the dynamic inventory management model of Ait-Sahalia and Saglam (2014) and its optimal quoting policy. Section 1.3 introduces the empirical investigation. Section 1.4 presents the data. Section 1.5 proposes a measure to determine the returns in a context of asynchronous data and dynamic inventory management. Section 1.6 sets out and discusses the results. Section 1.7 presents robustness tests, and Section 1.8 concludes the paper.

1.2 Optimal Quoting Policy of Liquidity Provision

Ait-Sahalia and Saglam (2014) refer to two types of agents: low-frequency traders (LFTs), who use market orders only, and a sole HFMM who has exclusive access to the limit order book (LOB). The HFMM trades limit orders (LOs) only, and exhibits inventory aversion. The bid-ask spread is exogenous. This setup differs from the classical dynamic inventory models in that the strategic variable is whether or not to quote and not to change the HFMM's supply curve. The HFMM's revenue depends on the trade-off between the inflows from the bid-ask spread and the outflows from the inventory cost as depicted in the following equation:

$$E(\pi) = C/2 \sum_{t=1}^{\infty} e^{-Dsmo_t} I(l_{smo_t}^b = 1) + C/2 \sum_{t=1}^{\infty} e^{-Dbmo_t} I(l_{bmo_t}^a = 1) - \Gamma \int_0^{\infty} e^{-Dt} |x_t| dt, \quad (1.1)$$

where:

$E(\pi)$: Quoting policy expected reward.

C : Bid-ask spread.

D : Constant discount factor > 0 .

smo_t (bmo_t): Sell (buy) market order by LFTs at time t .

I : Indicator function.

b : HFMM bid limit order.

a : HFMM ask limit order.

$l_{smo_t|bmo_t}^{b|a}$: Equals 1 if the HFMM is quoting a bid (b) or an ask (a) limit order when a LFT sell (buy) market order arrives, 0 otherwise.

Γ : Inventory aversion coefficient.

x_t : Inventory position at time t .

The first term to the right of equation (1.1) is the discounted value of the HFMM's revenue ($\frac{c}{2}$) earned when an incoming LFT's sell market order hits the HFMM's limit order while he is bidding ($I(l_{smo_t}^b = 1)$). The second term is the discounted revenue associated with an incoming LFT's bid market order, and the third term is the discounted value of the HFMM's inventory costs over the period dt . To keep the model tractable, the HFMM always places his LOs at the best bid and/or ask price and does not issue orders larger than one contract.

Apart from observing the arrival of market orders, the HFMM receives a signal s about the likely side of the next incoming market order: $s \in \{1, -1\}$, where 1 predicts an incoming LFT's buy market order and -1 an incoming LFT's sell market order. P quantifies the informational quality of the HFMM's signal. It varies from 0.5 (no prior knowledge about the side of the next incoming LFT's market order) to 1.0 (perfect knowledge). In Ait-Sahalia and Saglam (2014) setup, the next event is either 1: the arrival of a signal with probability $\left(\frac{\mu/2}{\lambda+\mu}\right)$, μ being the arrival rate of a Poisson distribution of the HFMM's signals and λ the arrival rate of a Poisson distribution of the incoming LFTs' market orders; 2: the arrival of a market order in the direction of the last signal with probability $\frac{P\lambda}{\lambda+\mu}$; or 3: the arrival of a market order in the opposite direction of the last signal with probability $\frac{(1-P)\lambda}{\lambda+\mu}$. The value of market-making activities for any given event assuming an inventory position of x ($x \in \{\dots, -2, -1, 0, 1, 2, \dots\}$) and a sell signal (-1) is:

$$v(x, -1) = -\gamma|x| + \delta \left\{ \left(\frac{\mu/2}{\lambda + \mu} \right) (v(x, 1) + v(x, -1)) + \frac{P\lambda}{\lambda + \mu} \max \left(\frac{c}{2\delta} + v(x - 1, -1), v(x, -1) \right) + \frac{(1-P)\lambda}{\lambda + \mu} \max \left(\frac{c}{2\delta} + v(x + 1, -1), v(x, -1) \right) \right\}, \quad (1.2)$$

where:

$$\gamma = \frac{\Gamma}{\lambda + \mu + D}; \delta = \frac{\lambda + \mu}{\lambda + \mu + D}; c = \delta C;$$

Equation (1.2) quantifies the market-making value function. The first term to the right is the discounted inventory cost $(-\gamma|x|)$. The second term is the discounted value of the three possible events: the value of the arrival of a signal $\left(\left(\frac{\mu/2}{\lambda + \mu} \right) (v(x, 1) + v(x, -1)) \right)$, the value of the arrival of a market order in the direction of the signal $\left(\frac{P\lambda}{\lambda + \mu} \max \left(\frac{c}{2\delta} + v(x - 1, -1), v(x, -1) \right) \right)$, and the value of the arrival of a market order in the opposite direction of the signal $\left(\frac{(1-P)\lambda}{\lambda + \mu} \max \left(\frac{c}{2\delta} + v(x + 1, -1), v(x, -1) \right) \right)$.

Solving equation (1.2) by backward induction using the Hamilton-Jacobi-Belman optimality method leads to the optimization of the expected reward trade-off.

1.2.1 OQP determination

Theorem 1 of Ait-Sahalia and Saglam (2014) states that there is an optimal quoting policy of liquidity provision, based on the expected reward trade-off:

Theorem 1: The optimal quoting policy of the HFMM consists in quoting at the best bid and the best ask according to a threshold policy, i.e., there exists $L^* < 0 < U^* \leq |L^*|$, such that:

$$l^b(x, 1) = \begin{cases} 1 & \text{when } x < U^* \\ 0 & \text{when } x \geq U^* \end{cases} \quad l^a(x, 1) = \begin{cases} 1 & \text{when } x > L^* \\ 0 & \text{when } x \leq L^* \end{cases}$$

$$l^b(x, -1) = \begin{cases} 1 & \text{when } x < -L^* \\ 0 & \text{when } x \geq -L^* \end{cases} \quad l^a(x, -1) = \begin{cases} 1 & \text{when } x > -U^* \\ 0 & \text{when } x \leq -U^* \end{cases}$$

Theorem 1 can be interpreted as follows: Suppose the HFMM receives a “buy” signal ($s = 1$) while being long ($x > 1$). He is going to act upon it ($l^b = 1$) as long as his current inventory is not already too high ($x < U^*$). If ($x \geq U^*$), the HFMM will not quote because this could increase his long inventory position beyond the optimal threshold U^* . Symmetrically, if the HFMM receives a “sell” signal ($s = -1$), he will quote on the ask side ($l^a = 1$) as long as his inventory position is not already too short ($x > -U^*$).

An algorithm proposed by Ait-Sahalia and Saglam (2014) presented in the Appendix 1.1 allows us to determine the thresholds based on the expected reward trade-off and Theorem 1.

1.2.2 Emulation and OQP

To determine an optimal quoting policy of liquidity provision, the model requires six parameters: $D, \Gamma, \lambda, \mu, C, P$. All financial instruments use the same and constant parameters D , the discount rate, and Γ , the coefficient of inventory aversion.

The four remaining parameters depend on the idiosyncratic behaviors of the stocks. C is the observed bid-ask spread and λ , the observed arrival rate of marketable orders. We define μ , the HFMM’s arrival rate of signals, as the number of creations, updates, and cancellations at the LOB level 1. We constrain the HFMM to react to other market participants’ actions. He does not use any private information to modify the observed price discovery process and/or the bid-ask spread. The parameter P is fixed at 0.50.

For any given combination of the six input parameters, we obtain an ex-ante OQP of liquidity provision based on the algorithm described in the Appendix. The algorithm stipulates the sides (bid and/or ask) and respective quantities to quote, i.e. the thresholds.

1.3 Empirical investigation

We aim to provide an empirical investigation of the profitability and the impact on market quality of an individual high-frequency trader acting as a liquidity provider. The decision to quote or to trade, the timing and the management of positions are totally driven by our fully automated algorithm. Our approach is fundamentally different from the traditional

trading strategy approaches such as Fibonacci ratios, golden ratio, oscillators and pivot point strategies that try to forecast the future value of a financial instrument. Our method involves using the optimal quoting policy from Ait-Sahalia and Saglam (2014) as a kernel. The OQP is independent from the market states and does not require any prediction of prices.

Data mining and data snooping have been analyzed extensively (Wasserstein and Lazar (2016), Bailey, Borwein et al. (2015), Kim and Ji (2015), and Bailey, Borwein et al. (2014)). Multiple-testing increases the probability of a false discovery drastically because it takes on average as few as $1/\alpha$ independent iterations to produce a false discovery (Lopez De Prado (2015)). Our results are obtained following a single set of parameter values designed ex-ante and therefore do not imply any data mining or data snooping.

First, we assess the performance of the optimal quoting policy from Ait-Sahalia and Saglam (2014). However, we impose the closing of all positions by issuing market orders at the end of the day (EOD); the procedure is launched at the beginning of the last three minutes of trading. The appraisal thus represents the results of “pure” market-making as accurately as possible.

Second, we embed the OQP in a trading strategy that considers market imperfections: limit orders are not uniquely identified in our database. Usually we cannot know with certainty who holds time priority. We apply the worst-case scenario to the HFMM: time priority is given to the total quantity available at the best bid (ask), excluding the HFMM limit order, one microsecond (μS) before the arrival of a market order. In this way, we depart from Ait-Sahalia and Saglam (2014), who assume that the HFMM is the fastest trader.

In practice, trading firms monitor market conditions and integrate pauses in their algorithms (Kelejian and Mukerji (2016)). Events like the flash crash of May 2010 and the Knight Capital’s algorithm glitch of August 2012¹ prompted regulators such as the Commodity Futures Trading Commission (2013), the U.S. Securities and Exchange

¹ <http://www.bloomberg.com/bw/articles/2012-08-02/knight-shows-how-to-lose-440-million-in-30-minutes> (Last accessed on November 24, 2016).

Commission (2016), and tThe Government Office for Science London (2012)) to make the use of circuit-breakers mandatory. We enforce circuit-breakers by monitoring market conditions associated with three of the OQP's parameters: λ and μ have an upper bound corresponding to 95% of the ranges of values from the reference time interval of one minute. C , the bid-ask spread, has an upper bound of 99% of the reference range. When a parameter's value exceeds its upper bound, we cancel all the quotes on the stock and we send a marketable order to liquidate the position. Parameter values are reset to zero at the beginning of the next time interval. This induces regular quoting and trading activities. This behavior is in line with Chordia, Goyal et al. (2013), who note that market-makers are also liquidity takers in their regular activities.

Within the model of Ait-Sahalia and Saglam (2014), the quantity of each order is fixed at one lot. To relax the constraint imposed on profits, we generalize this concept by defining κ , a constant quantity. κ is similar in nature to the trading unit of an option contract, e.g. 100 stocks, and is defined as the maximum quantity from the five most frequently traded quantities of a given stock. The fragmentation of orders is a well-established concept (Almgren and Chriss (2000); Almgren (2003); Obizhaeva and Wang (2013); Markov (2014); Jingle and Phadnis (2013); among others). Choosing κ with the suggested methodology reflects the fact that market participants want to mitigate their impact on the price discovery process. This is supported by the statistics of Table 1.2, which demonstrate that 47.6% (41.2%) of all trades in the DAX (MDAX) do not consume the available quantity at level one.

Trading a quantity larger than 1 could cause a price impact if the available quantity at level one is insufficient to liquidate the HFMM's position. This would force the HFMM to walk into the limit order book. Empirical investigations of trading strategies are vulnerable to biases if they exhibit price impacts. This could be an indication of an undue influence on the price discovery process and it could lead real-time trading results to differ significantly from expectations. To avoid HFMM's market orders and the implied illiquidity cost transfer to other market participants that affect prices, the HFMM incurs market risk instead of walking up (down) the LOB. The HFMM trades up to the available quantity at level one and waits for the next incoming limit order(s) at that level in order

to fully liquidate his position if necessary. This amounts to controlling for the instantaneous price impact (Cont, Kukanov et al. (2014); Bouchaud, Farmer et al. (2009)). Estimating the permanent impact of market orders (Hautsch and Huang (2012); Huh (2014); Zhou (2012)) becomes unnecessary. We apply to κ the OQP thresholds associated with the contemporary model's parameters. In case of partial execution of a limit order, we cancel the order(s) and submit a new order(s) with the required adjusted quantity(ies). Time priority is amended accordingly.

Speed is important to gain time priority and to avoid being picked up (sniped) while displaying stale quotes, so we take into consideration the effect of latency on the trading results. We use the latency of 150 μ S. This value is representative of the time required by our infrastructure and our algorithm to receive, analyze, react to, and send new orders following the arrival of new information. We can compete on speed with the other co-location firms, and we are significantly faster than buy-side investors.

1.4 Data

The data come from Xetra, the fully electronic trading platform of the Frankfurt Stock Exchange. The raw dataset contains all events (deltas and snapshots) sent through the Enhanced Broadcast System, a data feed used by high-frequency traders. Deltas track all possible events in the LOB whereas snapshots convey information about the state of a given LOB at a specific time. Xetra Parser, developed by Bilodeau (2013), is used to reconstruct the real-time order book sequence using Xetra protocol and Enhanced Broadcast. Liquidity is provided by market participants posting limit orders in the LOB. The stocks of our sample have LOB with twenty levels on both sides of the market. The state of the LOB and the arrival of marketable orders (the trades) can be observed by the subscribers to the data feed. Time stamps are in μ S, trading is anonymous and specific order identification is nonexistent.

Our data set consists of sixty stocks from the DAX index family: thirty stocks in each of the DAX and the MDAX. DAX indexes are indicators for the German equity market. The DAX characterizes the blue chip segment. Its components are the largest and most actively traded German companies. The MDAX is composed of mid-capitalization issues

from traditional sectors, excluding technology, that rank immediately below the DAX stocks. The ultra-high-frequency data cover six months from February 1, 2013 to July 31, 2013. The sample covers different market phases as depicted by Figure 1.1: a trading range for the entire month of February, a bull trend from the last week of April to mid-May, 2 bear trends (mid-March to mid-April and the last week of May to the last week of June) and high volatility periods: the third week of April and the third week of June. Figure 1.1 displays the DAX daily chart for this period.

[insert Figure 1.1 here]

Table 1.1 presents the summary statistics from trading and LOB level one quoting activities. In Panel A, the DAX largely dominates the MDAX with a market value traded of €398.5b (92.57% of total activities) and 17.6 m of transactions (79.47%). Panel B exhibits even stronger statistics for the DAX. Quoting based exclusively on level one activity overwhelms trading as depicted by the ratio of the number of updates to the number of trades (# UTDs/# trades) that is higher than 10 for both indexes. This ratio is followed by the SEC (MIDAS, Security and Exchange Commission at <http://www.sec.gov/marketstructure/midas.html>) to monitor high-frequency trading activities.

[insert Table 1.1 here]

Table 1.2 illustrates the price discovery process. Price impact minimization is the dominant trading phenomenon, with 47.6% (DAX) and 41.2% (MDAX) of all trades executed at the last tick price. This includes combinations (0,0), (+,0), and (-,0). Aggressive orders (+,+ and -,-) induce positive autocorrelations in the price discovery process. They represent 14.3% (DAX) and 18.4% (MDAX) of all price moves, less than bid-ask bounce trades (+,- and -,+), which are respectively 18.1% (DAX) and 21.0% (MDAX).

[insert Table 1.2 here]

Table 1.3 exhibits the two-way classification for the thirty stocks from the DAX index. The directional trading maxima (+,+ and -,-) are respectively 10.72% and 11.04% for the

unique identifier (isix) 1634, and the other aggregated results are representative of all stocks.

[insert Table 1.3 here]

Results for the thirty stocks from the MDAX are not qualitatively different. They are not presented due to space considerations, but are available upon request.

Johnson, Zhao et al. (2013) proposed the concept of ultrafast extreme events (UEEs). UEEs can shed light on the price discovery process and the instabilities of financial markets, and help one appraise the HFMM's risk exposure, and stress-test trading algorithms. We define UEEs as an occurrence of a stock price ticking down (up) at least five times before ticking up (down), having a price change of at least 0.5% within a duration of 1500 milliseconds. We can interpret UEEs as surges for up ticks and mini crashes for down ticks. As depicted in Table 1.4, three hundred and thirty-nine UEEs have been observed (85 in the DAX and 254 in the MDAX) during the 125 trading days of our sample (2.7 events on average per day). The number of events is significantly higher in the MDAX. This is consistent with the differences in liquidity and trading interest exhibited in Table 1.2. The difference in trading intensity between the DAX and the MDAX is also reflected in the higher average repetitions (7.918 vs 6.519).

[insert Table 1.4 here]

Figure 1.2 shows the number of UEE occurrences per day. Extreme events happened in 102 out of 125 trading days (84.30%). UEEs have occurred on the DAX (MDAX) during 40 (94) days. The higher number of daily UEE occurrences in the MDAX reflects its thinner trading and its shallower depth of LOB level one compared to the DAX. Spikes in the number of UEEs do not happen simultaneously in both indexes. This suggests that their causes are idiosyncratic rather than systematic.

[insert Figure 1.2 here]

Figure 1.3 displays the number of UEE occurrences per minute, considering the 510 minutes of trading on regular days. UEEs exhibit a tendency to occur around the open and

the close of the day as the documented smile in trading volume (Hanif and Smith (2012), Madhavan (2002)). UEEs can result from induced uncertainty by the market model that imposes long lasting suspension of trading.

[insert Figure 1.3 here]

The assumption of independent arrivals of HFMM signals and LFTs market orders is implicit in the Poisson distributions used by Ait-Sahalia and Saglam (2014). The HFMM decision follows the arrival of new information (LFTs marketable orders or signals), so we test the independence assumption on the aggregated information. Table 1.5 shows that one cannot reject this assumption for any stock in our sample.

[insert Table 1.5 here]

1.5 TVWAR: a time- and volume-weighted average return

Data emulation replicates the stock behavior. The model adapts the OQP dynamically to the stock's states by tracking the parameters λ (the arrival rate of LFTs' market orders), μ (the arrival rate of HFMM's signals), and C (the bid-ask spread). This induces dynamic management of positions. To evaluate the HFMM's performance, we propose a measure based on realized PnL, which considers the impacts of leverage and asynchronous data. Both factors affect the holding period and the discrete time returns.

Equations (1.3) to (1.6) define the variables required to determine the holding period return of a sequence of E events. A profit (loss) is realized when an existing position (long or short) is unwound. The unwinding quantity comes from two sources: the HFMM's marketable orders due to risk management features and incoming market orders executed against the HFMM's LOs. The unwinding quantity refers to a traded quantity that partially or totally offsets a position.

The unwinding quantity is:

$$\begin{aligned}
& UQ_e \\
& = \begin{cases} \text{if } Pos_{e-1} > 0 \text{ and } Pos_{e-1} > Pos_e & -\min(Pos_{e-1} - Pos_e, Pos_{e-1}) \\ \text{if } Pos_{e-1} < 0 \text{ and } Pos_{e-1} < Pos_e & -\max(Pos_{e-1} - Pos_e, Pos_{e-1}) \\ \text{else} & 0 \end{cases} \quad (1.3)
\end{aligned}$$

where:

UQ_e : unwinding quantity for event e , negative (positive) for buys (sells).

Pos_e : quantity long, short or flat for event e .

e : event number (an event = a trade).

$e \in [1, 2, \dots, E]$.

Maximum leverage ensues from two factors: the OQP, which depends on the parameters $(\Gamma, D, \lambda, \mu, C, P)$ of Ait-Sahalia and Saglam (2014) and κ , the reference quantity defined in Section 0. Effective leverage, with an upper bound equal to the maximum leverage, is influenced by speed (latency, time priority, and market-making competition), HFMM LOs, and incoming market orders (quantity and serial correlation).

The effective leverage value is:

$$\Phi_e = \left(\frac{UQ_e}{\kappa/2} \right), \quad (1.4)$$

where Φ_e is based on the required capital to trade κ shares considering standard margin requirements. An unwinding trade for the HFMM resulting in a partial execution of the reference quantity κ has a leverage value smaller than the leverage value of an unwinding trade that closes the HFMM position of 2κ . 2κ is possible when the OQP threshold is 2 and the HFMM carries the maximum inventory.

The holding period return of event e is:

$$r_e = \ln(P_e/P_{e-1}) \cdot \Phi, \quad (1.5)$$

where P_e is the trade price for event e .

Returns are directly impacted by the relative importance of the unwinding trades. All else being equal, there is a linear relationship between the leverage measure and the holding period return of an event.

The cumulated return over E events is:

$$r_E = \sum_{e=1}^E r_e. \quad (1.6)$$

Asynchronous events are the norm in microsecond trading environments. Significant differences exist in stocks' behavior due to their liquidity and depth, and to the trading interest. To facilitate comparison, we use discrete time intervals where the returns are time-weighted and volume-weighted within the interval. Equations (1.7) to (1.14) define the variables required to determine the discrete time return over D time intervals.

If a position overlaps two or more time intervals, the return is evenly spread out over the holding period. The number of time intervals in a trading day is equal to:

$$D = \lceil (\mu S_{eod} - \mu S_{bod}) / \mu S_{byInterval} \rceil, \quad (1.7)$$

where:

$d \in [1, 2, \dots, D]$.

μS_{eod} : time stamp of the end of the day in μS .

μS_{bod} : time stamp of the beginning of the day in μS .

$\mu S_{byInterval}$: number of μS in a one-time interval.

The reference time stamp at the beginning of period d is given by:

$$\mu SBegRef_d = \mu S_{bod} + (d - 1) \cdot \mu S_{byInterval}. \quad (1.8)$$

The reference time stamp at the end of period d equals:

$$\mu SEndRef_d = \min(\mu S_{bod} + d \cdot \mu S_{byInterval}, \mu S_{eod}). \quad (1.9)$$

The reference time stamp at the beginning of period d for event e solves:

$$\mu SBeg_{d,e} = \max(\mu S_{d,e}, \mu SBegRef_d), \quad (1.10)$$

where:

$\mu S_{d,e}$: time stamp of event e in time interval d .

The reference time stamp of period d for event e is bounded by the end of the time interval d , so the reference time stamp at the end of period d for event e is:

$$\mu SEnd_{d,e} = \min(\mu S_{d,e}, \mu SEndRef_d). \quad (1.11)$$

Our procedure allows us to consider both the leverage $\left(\frac{UQ_{d,e}}{\kappa/2}\right)$ and the holding period $\left(\frac{\mu SEnd_{d,e} - \mu SBeg_{d,e}}{\mu SByInterval}\right)$ of positions. The time- and volume-weighted return of event e during time interval d is:

$$r_{d,e} = \ln(P_{d,e}/P_{d,e-1}) \cdot \frac{UQ_{d,e}}{\kappa/2} \cdot \left(\frac{\mu SEnd_{d,e} - \mu SBeg_{d,e}}{\mu SByInterval}\right). \quad (1.12)$$

The return for time interval d is the sum of the time- and volume-weighted returns of the E events of period d :

$$r_d = \sum_{e=1}^E r_{d,e}. \quad (1.13)$$

The cumulated return over D time intervals adjusted for leverage and holding periods equals:

$$r_D = \sum_{d=1}^D r_d. \quad (1.14)$$

1.6 Results

We emulate the trading and quoting activities of an HFMM under two scenarios: Section 1.6.1 implements the OQP designed by Ait-Sahalia and Saglam (2014) dynamic inventory model. In contrast, we close all positions by issuing market orders before the end of the day. In that way, the results are as representative as possible of “pure” market-making activities. Whereas Ait-Sahalia and Saglam (2014) show that an HFMM who holds private information can exploit it to his advantage, we do not consider this opportunity because we do not want to interfere in the price discovery process. A latency of 150 microseconds is applied to consider the cycle of reception, analysis and response from our infrastructure.

The worst-case scenario is applied to the HFMM time priority as defined in Section 1.6.2. In Section 0, we relax the constraint on the quantity of each order fixed at 1 by using κ . The HFMM does not optimize on the quantity in each trade because κ is constant. The management of positions considers the dynamic nature of the OQP, the trading intensity and the liquidity needs of the market participants. Circuit-breakers are implemented and are based on the monitoring of market characteristics summarized by the arrival rates of new information (trades and quotes) and the behavior of the bid-ask spread. They involve the use of market orders. We avoid price impacts by restricting the quantity of market orders to the available quantity at level one. This forces the HFMM to incur market risk instead of transferring the liquidity risk to the other market participants by walking into the LOB. We analyze the impact of circuit-breakers on profitability and market quality. Profitability measures are from the Profit and Loss (PnL) report, which is calculated from all HFMM orders (limit and market). The approach to assess the PnL meets the requirements of the Basel Committee on Banking Supervision (2013).

1.6.1 Ait-Sahalia and Saglam (2014) model

Figure 1.4 displays the aggregated cumulative Profit and loss from the quoting and trading activities based on Ait-Sahalia and Saglam (2014) dynamic inventory model. Both indexes exhibit an upward trend over the entire period without significant drawdowns. The OQP allows the HFMM to extract a constant annuity from both indexes.

[insert Figure 1.4 here]

Daily and intraday statistics are presented in Table 1.6. Total profits for the six-month period are 3,412 € (2,999 €) in the DAX (MDAX). Total profits are 14.8% higher in the DAX than in the MDAX. This is the result of a higher number of trades, 181,594 vs 86,199 for the DAX vs. the MDAX combined with a tiny average profit per trade of 0.019 € vs 0.034 €. This is characteristic of high-frequency trading (Jones (2013)). Total profits are insufficient to maintain the infrastructure costs required by market-making activities.

[insert Table 1.6 here]

1.6.2 Trading strategy

The dynamic management of the OQP thresholds coupled with the use of κ based on the idiosyncratic characteristics of the stocks expose the HFMM's LOs to partial executions. Table 1.7 illustrates the way partial execution of the HFMM's limit orders against marketable orders are handled. The example comes from the emulation of Deutsche Bank data from February 1, 2013. It illustrates the way partial executions of the HFMM's limit orders against marketable orders are handled. It works as follow: One μ S before054552, the HFMM is short 500 shares. At054552, a bid limit order creation for 1,000 shares at 42.910 is sent to the Exchange. We identify this order with the internal id 41. Internal ids refer to emulated orders from the trading strategy. At139361, an incoming market order hits the HFMM bid for a quantity of 169 (internal id 7).

[insert Table 1.7 here]

The HFMM position is short 331 shares. This trade is immediately followed by the cancellation of the LO with id 41, a creation (id 42) of a bid LO at 42.910 for 831 shares and a creation of an ask LO at 42.950 for 169 shares. Considering the HFMM's position (short 331) and his LO quantity of 831 (169) on the bid (ask), a full execution of one of his LOs will lead to the HFMM's κ of 500 shares for Deutsche Bank. This quantity is the maximum of the five most traded quantities of Deutsche Bank as defined in Section 1.3.

The profitability of the strategy is displayed in Table 1.8. Panel A indicates that total profits are strongly positive. The average daily profits are economically significant, the standard deviations low, and no daily loss has occurred over the 125-day period even if the sample includes different kinds of market moods, including some high stress periods (See Figure 1.1).

[insert Table 1.8 here]

Panel B shows an average profit per trade of 2.32€ (1.73€) for the DAX (MDAX). When coupled with the total number of trades of Panel C, we obtain results typical of high-frequency trading: a high number of trades (1.1m for the DAX and 355.5k for the MDAX) paired with a small profit per trade. This highlights the fact that the HFMM is not a

directional trader. Using hard information, he benefits from the bid-ask bounces and the variability of bid-ask spreads due to varying liquidity levels without assuming the use of valuable private information. The lowest part of Panel C displays the distribution of trade profit per transaction. Whereas flat trades represent 20.53% (15.75%) of the 1.1 million (350k) trades, both distributions are skewed to the right. An HFMM acting as a designated sponsor and using the strategy has his transaction fees waived because he fulfills the Deutsche Boerse (2015) requirements.

The HFMM is a liquidity provider when he behaves as a market maker and issues market orders (consumes liquidity) for risk management purposes. Table 1.9 divides the HFMM's trade origins into three categories: the execution of the HFMM's LOs against incoming LFTs' marketable orders (LOB); the intraday liquidation of the HFMM's positions due to circuit-breakers (C.B.); and the closing of positions at the end of the day (O/N). This disentangles the liquidity-providing activities (LOB) from the liquidity-consuming ones (circuit-breakers and closing position at the end-of-day).

More than 94% (89%) of trades in the DAX (MDAX) originate from LOs. This confirms the HFMM's role as a liquidity provider. The monitoring of market condition has triggered 54.4k (28.9k) market orders in the DAX (MDAX). Nevertheless, the average profit per trade in both indexes is more than twice that obtained by LOB activities. An explanation is linked to the bid-ask spread: as volatility increases with extreme market conditions, the bid-ask spread widens. This has a direct and positive impact on the profitability of liquidated positions. Closing positions at the end of the day required an average of 1.75 (2.0) trades in the DAX (MDAX).

[insert Table 1.9 here]

Circuit-breakers (CBs) are important components of the strategy. They have been activated a daily average of 14.5 (7.7) times per stock ($DAX: 54,462 \div 125 \div 30$). Looking at the DAX components, market orders triggered by the circuit-breakers represent 4.87% of total trades and they account for 9.94% of the total PnL. Results are even more striking when looking at the MDAX: 8.07% of the trades come from the circuit-breakers for a contribution exceeding 30% of the total PnL. Avoiding overnight

positions comes at a cost: respectively 135,334 € and 15,086 € for the DAX and MDAX components.

Traditionally, informed traders are associated with high-volume transactions: Blume, Easley et al. (1994) show that volume provides information on information quality. Wang (1994) finds that volume is positively correlated with absolute changes in prices and dividends. Chakravarty, Gulen et al. (2004) relate informed trading in both stock and option markets to trading volume. However, developments in market structure and technological advances led to evidence of dynamic use of limit order strategies by which traders manage their positions: Bloomfield, O'Hara et al. (2005) use experimental asset markets to analyze make-take decisions in an electronic market. They note that informed traders' aggressive orders are replaced by limit orders as prices move toward fundamental values. Hasbrouck and Saar (2009) find evidence consistent with the use of a dynamic limit order strategy by which traders manage their positions on INET.

To investigate the impact of informed traders on HFMM's performance in this context, we compare the returns using a constant leverage to the discrete time returns adjusted for leverage and asynchronous trades using Equations (1.7) to (1.14). If large trades convey private information, the HFMM's performance adjusted for leverage should decline to reflect the permanent impact on fundamental value of the private information. The results in Table 1.10 mitigate this conclusion. In the DAX, the cumulated leveraged return is 99.9% higher than the cumulated constant return. This can be the result of mixed strategies using market and limit orders described by Easley, de Prado et al. (2016). This casts doubts on the quality of information obtained from trade volume alone. *A contrario*, the difference between the leveraged and constant return is -19.5% in the MDAX. The trade-off between market risk and liquidity risk in this segment could explain this phenomenon.

[insert Table 1.10 here]

Figure 1.5 displays the discrete time returns per time interval adjusted for leverage and asynchronous trades using Equations (1.7) to (1.14). Dynamic management of positions is well suited to benefit from UEEs because the best discrete returns are obtained at the opening and the closing of the trading session, where the majority of UEEs happen.

Moreover, time intervals 330-331 (327-329) and 420-423 (417-418) in the DAX (MDAX) coincide with a larger than usual number of UEE occurrences. The Deutsche Boerse market model, implying midday auctions, imposes a transfer of wealth between LFTs toward the HFMM as seen during period 243 (13h03) in the DAX: it exhibits a significant increase in average leveraged return for the HFMM. The same phenomenon is observed in the minute following the midday auction in the MDAX (period 246 at 13h06).

[insert Figure 1.5 here]

As depicted in Table 1.11, the HFMM would have had a positive impact on the market. He has traded €17.7b of market value, more than 90% of which comes from his limit orders.

[insert Table 1.11 here]

1.7 Robustness tests

1.7.1 Speed's impact on performance

Latency, the required time to receive, process, and react to new information, is considered crucial to the HFMM. Short latencies allow to limit being sniped on stale quotes, to aggregate new information rapidly and to gain access to market orders via time priority. Testing the effect of latency on performance is equivalent to quantify the impact of investments in technological infrastructures and softwares. It can serve as benchmark either to compare traders which differ solely by their speed or for capital budgeting decisions.

The treatment time of new information by the HFMM's strategy is 104 μ S. The lower bound of the latencies analyzed is 150 μ S to allow for order transmissions to the Exchange. We consider that the reference HFMM is using colocation facilities. To test for the impact of latency on the strategy, we use latencies of 150, 500, 1 500, 5 000, and 10 000 μ S. For the sixty stocks over the six-month period, we emulate real-time trading to obtain all limit and market orders. We calculate intraday and daily PnL. Finally, we aggregate the statistics by index. Results are presented in Table 1.12.

[insert Table 1.12 here]

For the DAX, decreasing latency does not influence the relative risk: $\sigma(\pi)/\pi$ is constant throughout all level of latencies. The investment in colocation services is fully justified by the augmentation of € 342 000 in total profits generated by the strategy between 10 000 μ S and 150 μ S latencies².

1.7.2 Strategy's features

The OQP from Ait-Sahalia and Saglam (2014) is the core of the HFMM's quoting decisions. It allows the obtaining of positive cash flows throughout the analyzed samples (ref. Section 1.6.1). However, their model of dynamic inventory management uses a quantity of one. This imposes a constraint on the profitability which is insufficient to cover the infrastructure costs required by the market-making activities. We relax this constraint using the kappa concept defined in Section 1.3. Furthermore, we analyze the impact of circuit-breakers and end-of-day closing of positions on performance. Table 1.13 presents the results obtained by emulating the HFMM's market-making activities for the thirty DAX's components during February to July 2013.

[insert Table 1.13 here]

Table 1.13 columns differ by the algorithm's active features. The Setup 4 performance, which includes kappa, the circuit breakers, and the closing position procedure, is the one presented throughout the paper. It is the one with maximum profit and minimum volatility. Allowing overnight positions (Setup 2) induces a marginal decrease in profits while volatility increases marginally. Intraday performances, as depicted by the ratios of average win/loss and the number of win/loss, confirms that Setup 4 and Setup 2 exhibit a similar behavior. Circuit breakers play a crucial role. Profitability almost cuts in half when this feature is deactivated (Setup 1 and Setup 3). This highlights the importance to react when market conditions changes as the bulk of order flow from circuit-breakers are cancellation

² 10 Gbits/s connections are available in data center in Frankfurt/Main, Germany for a monthly fee of € 4 500 ref: Deutsche Boerse (2015). "Price List for the Utilization of the Exchange EDP of FWB Frankfurt Stock Exchange and of the EDP XONTRO."

of limit orders standing at the LOB's level 1. Results are qualitatively the same when analyzing the thirty components of the MDAX.

1.8 Conclusion

We have implemented the optimal quoting policy of liquidity provision of Ait-Sahalia and Saglam (2014) without the assumption of valuable private information. Profits in both DAX and MDAX over the six-month period exhibit an upward trend without significant drawdowns. The OQP allows the HFMM to extract a constant annuity from the market. However, total profits are insufficient to maintain the infrastructure required by market-making activities.

We have embedded the OQP in a trading strategy. The trading strategy avoids data mining and data snooping. It considers latency and partial executions of limit and market orders. Special care has been taken to eliminate the price impact linked to the HFMM's trading and quoting activities. Circuit-breakers have been implemented in response to regulators' concerns.

The viability of the strategy has been established using an extensive dataset including sixty stocks in two market segments. It covers a six-month period where the market has encompassed drastically different phases. The strategy exhibits outstanding characteristics when risk and profitability are considered. Market-making activities in both indexes led to 3.45 million € in profit for the six-month period. This is realized through 1.5 million trades, and no daily loss is incurred for either of the indexes. These results are the lower bound of the potential HFMM's performance considering: 1) the worst-case scenario applied to his time priority, 2) the way partial executions are handled (loss of time priority, delay to cancel and re-enter quotes), 3) the constraint to incur market risk when using market orders implying quantities exceeding the available ones at level 1, 4) no informational advantage, 5) no maker-fee revenues, and 6) no valuable private information. We have disentangled the liquidity-providing role from liquidity-consuming activities. Whereas the core of the profits comes from quoting activities, the implementation of circuit breakers adds to total profits.

Ultrafast extreme events have been documented; they occur regularly. They are linked to idiosyncratic characteristics and do not exhibit systematic behavior. In a context where price impact minimization is a major concern for traders, the prevalence of UEEs deserves further research. These high-frequency events could be associated with elusive liquidity, predatory behaviors, algorithm glitches, and aggregation of information.

Because partial executions of orders (limit and market) are possible, we have proposed a methodology to determine the returns that simultaneously take into consideration the varying leverage and the asynchronous nature of high-frequency trading. This procedure lets one quantify the impact of high volume trades often attributed to informed traders. The effect of high volume trades on the HFMM's performance is inconclusive. Aggregated results show that the HFMM's performance increases in the DAX and decreases in the MDAX. The HFMM behaves like a constant liquidity provider and has a positive effect on market quality.

As expected, latency affects performance. Investments in infrastructure and software are warranted by the increase in profitability and the HFMM can exploit his speed's advantage to economically significant levels. Empirical research must address market imperfections as they have considerable impact on both risk and profitability. Implementing circuit-breakers is crucial to the economic viability of the HFMM.

Ait-Sahalia and Sağlam (2016) have recently published an extension to the model analyzed in this paper. They endogenize the bid-ask spread which is a function of the HFMM's quoting decisions. These decisions are driven by a signal about the likely type of trader (patient or impatient) who will send the next incoming marketable order. To test adequately this new setup, one has to quote and trade actively in live markets as it involves both liquidity provision and supply curve decisions which are modifying the state of the LOB and the price discovery process. No academic financial laboratory is available around the world (Lopez De Prado (2015)) to realize that kind of test.

References

- Ait-Sahalia, Y. and M. Saglam (2014). "High Frequency Traders : Taking Advantage of Speed." SSRN Electronic Journal.
- Ait-Sahalia, Y. and M. Sağlam (2016). "High Frequency Market Making." SSRN Electronic Journal.
- Almgren, R. (2003). "Optimal Execution with Nonlinear Impact Functions and Trading-Enhanced Risk." Applied Mathematical Finance **10**(1): 1-18.
- Almgren, R. and N. Chriss (2000). "Optimal Execution of Portfolio Transactions." Journal of Risk **3**(2): 5-39.
- Anand, A. and K. Venkataraman (2013). "Should Exchanges Impose Market Maker Obligations?" SSRN Electronic Journal.
- Bailey, D. H., et al. (2014). "Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-of-Sample Performance." Notices of the American Mathematical Society **61**(5): 458-471.
- Bailey, D. H., et al. (2015). "The Probability of Backtest Overfitting." SSRN Electronic Journal.
- Basel Committee on Banking Supervision (2013). "Fundamental Review of the Trading Book: A Revised Market Risk Framework." Research Report.
- Biais, B., et al. (2015). "Equilibrium Fast Trading." Journal of Financial Economics **116**(2): 292-313.
- Bilodeau, Y. (2013). "Xetraparser [computer software]."
- Bloomfield, R., et al. (2005). "The "Make or Take" Decision in an Electronic Market: Evidence on the Evolution of Liquidity." Journal of Financial Economics **75**(1): 165-199.
- Blume, L., et al. (1994). "Market Statistics and Technical Analysis: The Role of Volume." The Journal of Finance **49**(1): 153-181.
- Bouchaud, J. P., et al. (2009). "How Markets Slowly Digest Changes in Supply and Demand." In Handbook of Financial Markets: Dynamics and Evolution, edited by T. Hens and K.R. Schenk-Hoppé, pp. 57–160.
- Brogaard, J., et al. (2016). "High-Frequency Trading and Extreme Price Movements." SSRN Electronic Journal.
- Chakravarty, S., et al. (2004). "Informed Trading in Stock and Option Markets." The Journal of Finance **59**(3): 1235-1258.
- Chordia, T., et al. (2013). "High-frequency Trading." Journal of Financial Markets **16**(4): 637-645.
- Commodity Futures Trading Commission (2013). "Concept Release on Risk Controls and System Safeguards for Automated Trading Environments." Research Report.

- Cont, R., et al. (2014). "The Price Impact of Order Book Events." Journal of Financial Econometrics **12**(1): 47-88.
- Deutsche Boerse (2015). "Price List for the Utilization of the Exchange EDP of FWB Frankfurt Stock Exchange and of the EDP XONTRO."
- Duffie, D. (2010). "Presidential Address: Asset Price Dynamics with Slow-Moving Capital." The Journal of Finance **65**(4): 1237-1267.
- Easley, D., et al. (2016). "Discerning Information from Trade Data." Journal of Financial Economics **120**(2): 269-285.
- Foucault, T., et al. (2013). "Liquidity Cycles and Make/Take Fees in Electronic Markets." Journal of Finance **68**(1): 299-341.
- Glosten, L. R. and P. R. Milgrom (1985). "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." Journal of Financial Economics **14**(1): 71-100.
- Grossman, S. J. and J. E. Stiglitz (1980). "On the Impossibility of Informationally Efficient Markets." American Economic Review **70**(3): 393-408.
- Hanif, A. and R. E. Smith (2012). "Algorithmic, Electronic, and Automated Trading." The Journal of Trading **7**(4): 78-87.
- Hasbrouck, J. and G. Saar (2009). "Technology and Liquidity Provision: The Blurring of Traditional Definitions." Journal of Financial Markets **12**(2): 143-172.
- Hautsch, N. and R. Huang (2012). "The Market Impact of a Limit Order." Journal of Economic Dynamics & Control **36**: 501-522.
- Huh, S.-W. (2014). "Price Impact and Asset Pricing." Journal of Financial Markets **19**: 1-38.
- Jingle, L. I. U. and K. Phadnis (2013). "Optimal Trading Algorithm Selection and Utilization: Traders' Consensus versus Reality." Journal of Trading **8**(4): 9-19.
- Johnson, N., et al. (2013). "Abrupt Rise of New Machine Ecology Beyond Human Response Time." Scientific reports **3**: 2627.
- Jones, C. M. (2013). "What Do We Know About High-frequency Trading?" Working Paper, Columbia University.
- Kelejian, H. H. and P. Mukerji (2016). "Does High Frequency Algorithmic Trading Matter for Non-AT Investors?" Research in International Business and Finance **37**: 78-92.
- Kim, J. H. and P. I. Ji (2015). "Significance Testing in Empirical Finance: A Critical Review and Assessment." Journal of Empirical Finance **34**: 1-14.
- Kyle, A. S. (1985). "Continuous Auctions and Insider Trading." Econometrica **53**(6): 1315-1335.

- Lopez De Prado, M. (2015). "The Future of Empirical Finance." Journal of Portfolio Management **41**(4): 140-144.
- Madhavan, A. (2002). "VWAP Strategies." Institutional Investor Journals **1**: 32-39.
- Malinova, K. and A. Park (2015). "Subsidizing Liquidity: The Impact of Make/Take Fees on Market Quality." The Journal of Finance **70**(2): 509-536.
- Markov, V. (2014). "Constant Impact Strategy." Journal of Trading **9**(3): 26-33.
- Menkveld, A. J. (2013). "High Frequency Trading and the New-Market Makers." Journal of Financial Markets **16**(4): 712-740.
- Obizhaeva, A. A. and J. Wang (2013). "Optimal Trading Strategy and Supply/Demand Dynamics." Journal of Financial Markets **16**(1): 1-32.
- Pagnotta, E. and T. Philippon (2015). "Competing on Speed." SSRN Electronic Journal.
- Popper, N. (2012). "High-speed Trading no Longer Hurling Forward." The New York Times (October 14).
- Roll, R. (1984). "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market." The Journal of Finance **39**(4): 1127-1139.
- Serbera, J.-P. and P. Paumard (2016). "The Fall of High-frequency Trading: A Survey of Competition and Profits." Research in International Business and Finance **36**: 271-287.
- The Government Office for Science London (2012). "EIA4: Circuit Breakers." Research Report.
- U.S. Securities and Exchange Commission (2016). "Investor Bulletin: Measures to Address Market Volatility, Update."
- Wang, J. (1994). "A Model of Competitive Stock Trading Volume." Journal of Political Economy **102**(1): 127-168.
- Wasserstein, R. L. and N. A. Lazar (2016). "The ASA's Statement on p-Values: Context, Process, and Purpose." The American Statistician **70**(2): 129-133.
- Zhou, W.-X. (2012). "Universal Price Impact Functions of Individual Trades in an Order-driven Market." Quantitative Finance **12**(8): 1253-1263.

Appendix 1.1: Threshold calculation: an efficient algorithm

Output: L^*, U^*

Initialize $L = 1$ and $flag = 0$

While $flag = 0$ do

$U \leftarrow 1$;

While $U \leq -L$ do

Solve for $v(L-1, 1), v(L, 1), \dots, v(-L, 1), v(-L+1, 1)$;

if

$v(L, 1) - v(L-1, 1) > C/2\delta, v(L+1, 1) - v(L, 1) \leq C/2\delta, v(U, 1) - v(U+1, 1) > C/2\delta, v(U-1, 1) - v(U, 1) \leq C/2\delta$

Then

$flag \leftarrow 1, L^* \leftarrow L$ and $U^* \leftarrow U$;

Break;

$U \leftarrow U + 1$;

$L \leftarrow L - 1$;

Where:

L : Lower bound threshold.

U : Upper bound threshold.

$v(L|U, s)$: Value function at threshold L or U , having a signal s
 $s \in \{1, -1\}$. 1 denotes a buy signal and -1 denotes a sell signal.

$c = \delta C$.

C : Bid-ask spread.

$\delta = \frac{\lambda + \mu}{\lambda + \mu + D}$.

λ : Arrival rate of LFTs market orders

μ : Arrival rate of HFMM signals s .

D : Constant discount factor > 0 .

Γ : Inventory aversion coefficient.

Table 1.1 Market summary statistics, trades and LOB1

Panel A					
Trades	Market Value (MV)	% MV	# Trades	% Trades	
DAX (30)	€ 398,504,790,578	92.57%	17,637,381	79.47%	
MDAX (30)	€ 31,986,673,636	7.43%	4,557,183	20.53%	
Total	€ 430,491,464,214	100.00%	22,194,564	100.00%	
Panel B					
LOB, level 1	Market Value (MV)	% MV	# UTD	% UTD	# UTD/ # Trades
DAX (30)	€ 11,038,241,222,840	95.48%	207,225,811	80.96%	11.75
MDAX (30)	€ 522,821,561,308	4.52%	48,740,327	19.04%	10.70
	€ 11,561,062,784,148	100.00%	255,966,138	100.00%	11.53

The data span is from February 2 to July 30, 2013. #UTD/#Trades are used to monitor high frequency trading activities, as for MIDAS, Security and Exchange Commission (SEC) at <http://www.sec.gov/marketstructure/midas.html>.

Table 1.2 Two-way classification of price movements in consecutive intraday trades: summary (,000)

		+,+	+,0	+, -	0, +	0, 0	0, -	- , +	- , 0	- , -	Total
DAX	Occ	.	1,253.8	1,776.1	1,599.7	1,781.8	4,824.6	1,770.0	1,594.0	1,775.7	17,637.4
	%		7.1%	10.1%	9.1%	10.1%	27.4%	10.0%	9.0%	10.1%	100.0%
MDA	Occ	.	418.0	445.0	478.6	443.9	988.7	442.3	479.6	441.2	4,557.2
X	%		9.2%	9.8%	10.5%	9.7%	21.7%	9.7%	10.5%	9.7%	100.0%

Price movements are classified into “up” (+), “unchanged” (0), and “down” (-). Price moves are represented by x,y where x is the ith-1 move and y the ith move. % is the relative occurrence of the column’s price movement.

Table 1.3 Two-way classification of price movements in consecutive intraday trades DAX (%)

Isix	+,+	+,0	+,-	0,+	0,0	0,-	-,+	-,0	-,-
22	8.36%	10.12%	9.63%	10.22%	23.33%	10.12%	9.53%	10.22%	8.47%
24	8.90%	10.64%	8.81%	10.63%	22.20%	10.58%	8.82%	10.57%	8.86%
32	9.51%	9.78%	9.47%	9.70%	23.74%	9.56%	9.55%	9.48%	9.23%
49	7.22%	10.59%	8.80%	10.59%	24.94%	10.84%	8.79%	10.84%	7.38%
58	8.61%	10.44%	9.40%	10.52%	22.47%	10.32%	9.32%	10.40%	8.53%
60	9.72%	10.39%	10.17%	10.32%	19.36%	10.13%	10.24%	10.05%	9.63%
80	3.23%	10.13%	6.04%	10.36%	40.89%	10.01%	5.80%	10.25%	3.28%
85	3.92%	10.41%	5.97%	10.33%	38.66%	10.36%	6.05%	10.28%	4.02%
106	5.94%	10.44%	8.16%	10.59%	29.99%	10.43%	8.02%	10.58%	5.85%
130	4.06%	10.10%	6.98%	10.38%	37.63%	9.88%	6.70%	10.16%	4.11%
138	2.14%	10.29%	7.66%	10.03%	39.89%	9.99%	7.93%	9.73%	2.34%
143	5.01%	10.64%	6.83%	10.44%	33.72%	10.74%	7.03%	10.54%	5.06%
146	1.97%	8.84%	5.42%	8.98%	49.76%	8.80%	5.28%	8.94%	2.02%
151	3.45%	10.01%	6.25%	10.28%	40.55%	9.90%	5.97%	10.18%	3.41%
266	7.55%	9.96%	10.34%	10.16%	24.34%	9.79%	10.14%	9.99%	7.74%
829	8.96%	10.00%	10.11%	9.97%	22.06%	9.98%	10.14%	9.94%	8.83%
1634	10.72%	9.95%	10.78%	10.01%	17.47%	9.63%	10.72%	9.69%	11.04%
2451	8.99%	9.65%	10.76%	9.51%	21.45%	9.79%	10.91%	9.64%	9.30%
2481	5.43%	10.45%	8.01%	10.49%	31.47%	10.37%	7.97%	10.41%	5.39%
2807	7.30%	10.42%	9.36%	10.22%	25.09%	10.49%	9.56%	10.29%	7.28%
2841	8.46%	9.92%	10.34%	9.89%	22.13%	10.14%	10.38%	10.10%	8.64%
3446	8.71%	10.29%	9.63%	10.24%	22.51%	10.12%	9.67%	10.08%	8.74%
3679	7.33%	9.79%	9.70%	9.75%	26.56%	9.92%	9.74%	9.87%	7.35%
3744	2.65%	9.60%	6.44%	9.78%	43.44%	9.46%	6.26%	9.64%	2.74%
4423	8.84%	10.13%	10.25%	10.30%	21.33%	10.03%	10.08%	10.20%	8.84%
5830	8.86%	10.19%	10.21%	10.16%	21.35%	10.15%	10.24%	10.12%	8.72%
8669	9.80%	9.92%	10.53%	9.87%	20.13%	9.83%	10.57%	9.78%	9.57%
9633	7.72%	10.12%	9.98%	10.33%	23.91%	10.10%	9.78%	10.30%	7.77%
11814	4.58%	10.34%	7.52%	10.40%	35.07%	10.00%	7.46%	10.07%	4.55%
16753	7.21%	10.24%	9.24%	10.34%	26.09%	10.12%	9.14%	10.22%	7.40%

Isix: unique stock identifier. Price movements are classified into “up” (+), “unchanged” (0), and “down” (-). Price moves are represented by x,y where x is the i^{th-1} move and y the i^{th} move. % is the relative occurrence of the column’s price movement. Each row sums to 1. The data span is from February 2, 2013 to July 30, 2013.

Table 1.4 Ultrafast extreme events (UEEs) summary

	UEEs up				UEEs down				Total UEEs			
	# occ.	# stocks	avg. rep.	# days	# occ.	# stocks	avg. rep.	# days	# occ.	# stocks	avg. rep.	# days
DAX	33	18	8.333	23	52	23	7.654	32	85	26	7.918	40
MDA												
X	133	28	6.436	72	121	27	6.653	64	254	30	6.519	94
Total	166	46	6.813	82	173	50	6.954	76	339	56	6.885	102

UEEs up: surges in price; UEEs down: mini crashes in price; # occ: number of UEE occurrences; # stocks: number of stocks that experienced at least one UEE over the sample; avg. rep.: average number of successive tick up (tick down) by UEE; # days: number of days with at least one UEE.

Table 1.5 Signal and trade independence: chi-square tests

DAX				MDAX			
Isix	X^2	p_val	#/ μ S	isix	X^2	p_val	#/ μ S
22	203,937,984	0.496	999,696	39	15,329,466	0.499	851,637
24	160,914,996	0.496	981,189	54	13,609,928	0.499	800,584
32	173,733,875	0.496	992,765	63	14,478,948	0.499	804,386
49	124,653,438	0.497	989,313	68	10,562,100	0.499	704,140
58	139,181,664	0.497	987,104	86	15,369,156	0.499	853,842
60	134,565,880	0.497	989,455	95	8,703,645	0.499	580,243
80	84,408,825	0.497	993,045	98	15,138,738	0.499	841,041
85	82,746,488	0.497	940,301	112	9,590,670	0.499	639,378
106	104,942,040	0.497	999,448	117	19,692,002	0.499	895,091
130	101,905,854	0.497	999,077	177	6,357,416	0.498	489,032
138	57,920,763	0.497	864,489	661	17,147,576	0.499	902,504
143	111,044,309	0.497	982,693	1131	10,887,405	0.499	725,827
146	144,863,120	0.497	999,056	1415	13,654,791	0.499	803,223
151	92,681,940	0.497	996,580	1429	14,507,892	0.499	805,994
266	89,374,050	0.497	993,045	1457	15,034,662	0.499	835,259
829	125,122,410	0.497	993,035	1468	7,207,956	0.498	514,854
1634	125,912,901	0.497	976,069	1566	16,050,003	0.499	844,737
2451	156,993,249	0.496	999,957	2323	15,455,538	0.499	858,641
2481	98,013,663	0.497	990,037	3290	5,190,090	0.498	346,006
2807	85,972,566	0.497	999,681	3849	11,741,355	0.499	782,757
2841	143,999,856	0.497	999,999	4035	13,745,027	0.499	808,531
3446	176,847,628	0.496	993,526	5566	10,716,976	0.499	669,811
3679	144,857,175	0.497	999,015	8650	11,551,908	0.499	679,524
3744	131,883,576	0.497	999,118	10658	10,201,716	0.499	728,694
4423	115,978,239	0.497	991,267	10938	13,980,222	0.499	822,366
5830	174,965,700	0.496	999,804	11426	9,910,173	0.499	762,321
8669	131,157,551	0.497	986,147	11475	13,561,136	0.499	847,571
9633	145,996,204	0.497	999,974	11607	7,284,465	0.498	485,631
11814	105,096,684	0.497	982,212	11644	7,226,336	0.499	555,872
16753	133,886,502	0.497	999,153	13469	16,029,576	0.499	890,532

Isix: unique stock identifier; X^2 : chi-square test value; p_val: p_value; #/ μ S: number of microseconds with at least one signal.

Table 1.6 Daily and intraday profitability – OQP

Panel A	Daily Statistics			
		DAX (30)		MDAX (30)
Total	€	3,412	€	2,999
Avg	€	27.30	€	23.99
Min	€	(33.49)	€	(29.62)
Max	€	103.84	€	125.54
Std. dev.	€	23.77	€	16.39
Days		125		125
Panel B	Intraday Statistics			
		DAX (30)		MDAX (30)
# trades		181,594		86,199
Avg π/trade	€	0.019	€	0.034
Min	€	(13.35)	€	(14.00)
Max	€	2.00	€	10.90

Table encompasses results from February 2 to July 30 2013; Total: Total profit; Avg: Average profit; Min: Minimum profit; Max: Maximum profit; Std. dev.: standard deviation of daily profit; Days: Number of trading days; Avg π /trade: Average profit per trade; # trades: Total number of executed trades by the HFMM.

Table 1.7 Orders and positions: an example

Isix: 2481

Date	Time	Type	Price	Q	Id	Pos
.	-500
20130201	32901054552	1	42.910	1000	41	-500
20130201	32901139361	3	42.910	169	7	-331
20130201	32901139361	2	42.910	-1000	41	-331
20130201	32901139361	1	42.910	831	42	-331
20130201	32901139361	1	42.925	-169	-103	-331
20130201	32907274679	2	42.910	-831	42	-331
20130201	32907274679	1	42.890	831	43	-331
20130201	32907276733	2	42.925	169	-103	-331

Type: 1 = HFMM new limit order, 2 = HFMM limit order cancellation, 3 = incoming LFT market order executed against an HFMM limit order; Time stamps are in microseconds; q = order quantity; id: internal reference to algorithm activity; pos: HFMM position, negative values representing short positions.

Table 1.8 Daily and intraday profitability – trading strategy

Panel A	Daily statistics			
	DAX (30)		MDAX (30)	
Total Profit	€	2,765,462	€	686,726
Avg	€	22,124	€	5,494
Min	€	7,396	€	1,035
Max	€	48,035	€	13,805
Std. Dev	€	8,314	€	2,349
No. obs		125	€	125
	Intraday statistics			
Panel B				
Avg π/trade	€	2.48	€	1.94
Panel C	Distribution of trades per profit			
	# trans	% total	# trans	% total
Total	1,113,352	100%	353,521	100.00%
<= -20	34,064	3.06%	17,227	4.87%
<-10 ; >= -20	51,135	4.59%	18,131	5.13%
<0 ; <= -10	268,015	24.07%	92,404	26.14%
0	228,603	20.53%	55,684	15.75%
>0 ; <= 10	333,731	29.98%	114,007	32.25%
> 10 ; <= 20	111,497	10.01%	28,545	8.07%
>20	86,307	7.75%	27,523	7.79%

Table encompasses results from February 2 to July 30 2013; Avg: Average daily profit; Min: Minimum daily profit; Max: Maximum daily profit; Std. dev.: standard deviation of daily profit; No. obs.: Number of trading days; Avg π /trade: Average profit per trade; Total: Total number of executed trades by the HFMM; <= -20, <-10 >-20, ..., <=20: bins of number of trades with profit <= -20, <-10 >-20, ..., <=20.

Table 1.9 HFMM's trade origins

		# trades	% trades	MV (000 €)	% MV	P&L	% PL
DAX	Total	1,117,499	100.00%	€ 16,021,276	100.00%	€ 2,765,462	100.00%
	LOB	1,056,471	94.54%	€ 15,311,005	95.57%	€ 2,625,985	94.96%
	C.B.	54,462	4.87%	€ 642,880	4.01%	€ 274,811	9.94%
	O/N	6,566	0.59%	€ 67,362	0.42%	€ (135,334)	-4.89%
MDAX	Total	357,599	100.00%	€ 1,696,227	100.00%	€ 686,726	100.00%
	LOB	321,175	89.81%	€ 1,575,293	92.87%	€ 493,181	71.82%
	C.B.	28,860	8.07%	€ 97,511	5.75%	€ 208,631	30.38%
	O/N	7,564	2.12%	€ 23,399	1.38%	€ (15,086)	-2.20%

trades: HFMM number of trades; MV (000€): € market value of HFMM trades (in thousands); P&L: Profit (loss); LOB: HFMM limit orders executed against incoming LFTs' market orders; C.B.: circuit-breakers (HFMM market orders due to real-time monitoring of market conditions); O/N: HFMM market orders to flatten position overnight.

Table 1.10 Impact of leverage on performance

DAX	Lev.	Constant		MDAX	Lev.	Constant	
Avg	878.8%	778.9%	99.9%	Avg	557.2%	576.7%	-19.5%
Isix	Cumul	Cumul	Diff	isix	Cumul	Cumul	Diff
22	968.4%	862.7%	105.7%	39	1034.4%	1003.9%	30.6%
24	408.1%	481.1%	-72.9%	54	172.6%	158.3%	14.3%
32	723.2%	855.8%	-132.6%	63	580.7%	542.8%	37.8%
49	585.3%	552.2%	33.1%	68	381.6%	357.8%	23.9%
58	690.5%	1053.0%	-362.5%	86	380.3%	488.2%	-107.9%
60	625.2%	742.6%	-117.4%	95	482.7%	231.6%	251.1%
80	751.3%	725.5%	25.8%	98	703.5%	813.0%	-109.6%
85	557.4%	413.5%	143.9%	112	549.4%	545.2%	4.1%
106	1327.3%	1191.0%	136.3%	117	719.9%	738.4%	-18.5%
130	1288.3%	1122.5%	165.8%	177	834.5%	522.9%	311.6%
138	756.1%	590.4%	165.7%	661	756.6%	698.8%	57.8%
143	712.4%	495.3%	217.1%	1131	686.9%	760.3%	-73.3%
146	1084.9%	786.9%	297.9%	1415	599.0%	644.5%	-45.5%
151	931.4%	731.3%	200.1%	1429	705.3%	653.2%	52.1%
266	327.0%	347.1%	-20.0%	1457	480.5%	487.4%	-6.9%
829	963.1%	636.5%	326.5%	1468	34.4%	47.3%	-12.9%
1634	520.9%	594.7%	-73.8%	1566	874.9%	754.1%	120.8%
2451	1357.5%	1462.5%	-105.0%	2323	479.2%	424.9%	54.3%
2481	96.7%	104.3%	-7.6%	3290	349.5%	582.0%	-232.4%
2807	963.2%	873.6%	89.6%	3849	853.2%	1210.0%	-356.8%
2841	1721.7%	1490.5%	231.2%	4035	239.5%	236.4%	3.1%
3446	829.4%	771.6%	57.8%	5566	700.3%	768.8%	-68.6%
3679	931.2%	716.5%	214.7%	8650	577.6%	616.9%	-39.2%
3744	1245.4%	834.7%	410.6%	10658	392.0%	288.6%	103.4%
4423	857.7%	818.4%	39.3%	10938	534.5%	434.5%	100.0%
5830	1171.5%	566.1%	605.4%	11426	704.9%	795.1%	-90.2%
8669	520.5%	617.9%	-97.4%	11475	967.6%	1492.3%	-524.6%
9633	1313.6%	1118.9%	194.7%	11607	298.0%	367.0%	-69.0%
11814	867.9%	753.3%	114.6%	11644	256.9%	260.1%	-3.2%
16753	1266.6%	1057.2%	209.4%	13469	385.2%	375.7%	9.5%

Lev. Cumul: cumulated return adjusted for leverage; Constant cumul: cumulated return with constant leverage.

Table 1.11 Participation in trades

		Realized	HFMM	HFMM/ Realized
DAX (30)	Market Value	€ 398,504,790,578	€ 16,021,291,071	4.02%
	# Trades	17,637,381	1,117,499	6.34%
MDAX (30)	Market Value	€ 31,986,673,636	€ 1,696,246,094	5.30%
	# Trades	4,557,183	357,599	7.85%

Realized column presents the summary statistics of realized market activities over the sample period (February 2 to July 30 2013). HFMM column presents what would have been the HFMM's activities over the same period.

Table 1.12 Impact of latency on performance

	Latency	10,000	5,000	2,500	1,000	500	150
DAX	π	€ 2,578,400	€ 2,784,300	€ 2,860,600	€ 2,898,200	€ 2,913,199	€ 2,765,462
	$\bar{\pi}$	€ 20,628	€ 22,274	€ 22,885	€ 23,816	€ 23,305	€ 22,124
	$\min(\pi)$	€ 8,182	€ 9,239	€ 8,799	€ 8,987	€ 8,952	€ 7,396
	$\sigma(\pi)$	€ 7,609	€ 8,115	€ 8,420	€ 8,582	€ 8,635	€ 8,314
	$\sigma(\pi)/\bar{\pi}$	0.37	0.36	0.37	0.36	0.37	0.37
MDAX	π	€ 609,760	€ 662,230	€ 691,710	€ 709,940	€ 714,640	€ 686,726
	$\bar{\pi}$	€ 4,878	€ 5,298	€ 5,534	€ 5,680	€ 5,717	€ 5,494
	$\min(\pi)$	€ 1,390	€ 1,724	€ 1,794	€ 1,783	€ 2,153	€ 1,035
	$\sigma(\pi)$	€ 1,715	€ 1,787	€ 1,850	€ 1,893	€ 1,906	€ 2,349
	$\sigma(\pi)/\bar{\pi}$	0.35	0.34	0.33	0.33	0.33	0.43

π : total profit; $\bar{\pi}$: average daily profit; $\sigma(\pi)$: standard deviation of daily profit; $\min(\pi)$: minimum daily profit.

Table 1.13 Impact from strategy's features

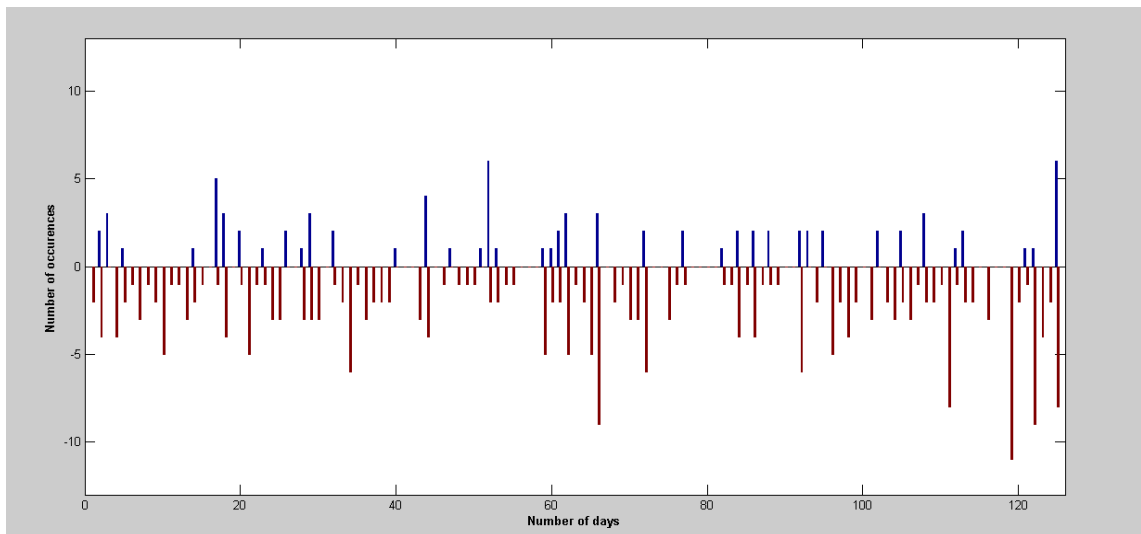
DAX (30)		Setup 1	Setup 2	Setup 3	Setup 4
daily	tot profit	1,421,867	2,722,256	1,435,176	2,765,462
	avg	11,375	21,778	11,481	22,124
	min	-10,921	6,400	3,488	7,396
	max	34,955	53,189	30,258	48,035
	std dev	9,063	8,439	5,360	8,314
trades	# of days	125	125	125	125
	avg	2.07	2.45	2.07	2.48
	min	-7,440	-7,440	-2,610	-1,160
	max	1,882	1,858	1,866	2,745
	std dev	42.77	29.90	20.40	17.17
	# trades	688,087	1,108,920	692,579	1,113,352
	<= -20	18,878	36,097	22,045	34,064
	<-10 ; >= -20	23,712	40,538	33,741	51,135
	<0 ; <= -10	134,850	213,615	177,390	268,015
	0	217,542	345,738	118,308	228,603
	>0 ; <= 10	183,213	289,639	217,977	333,731
	> 10 ; <= 20	53,408	88,207	71,408	111,497
	>20	56,484	95,086	51,710	86,307
	avg win	12.87	13.22	10.44	10.94
	avg loss	-13.25	-12.16	-9.12	-8.64
	avg w/l	0.97	1.09	1.15	1.27
	# win	293,105	472,932	341,095	531,535
	# loss	177,440	290,250	233,176	353,214
	# w/l	1.65	1.63	1.46	1.5

Setup 1: Circuit-breakers: off ; EOD liquidation: off ; Setup 2: Circuit-breakers: on ; EOD liquidation: off;
Setup 3: Circuit-breakers: off ; EOD liquidation: on ; Setup 4: Circuit-breakers: on ; EOD liquidation: on.

Figure 1. 1 DAX daily quotes - February to July 2013

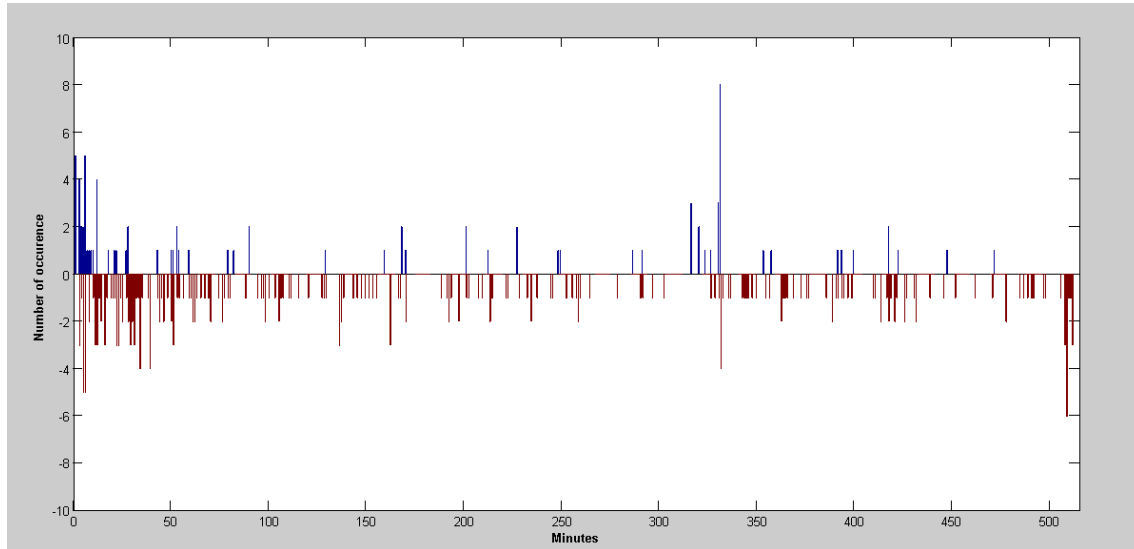


Figure 1. 2 Number of UEEs per day: DAX - MDAX



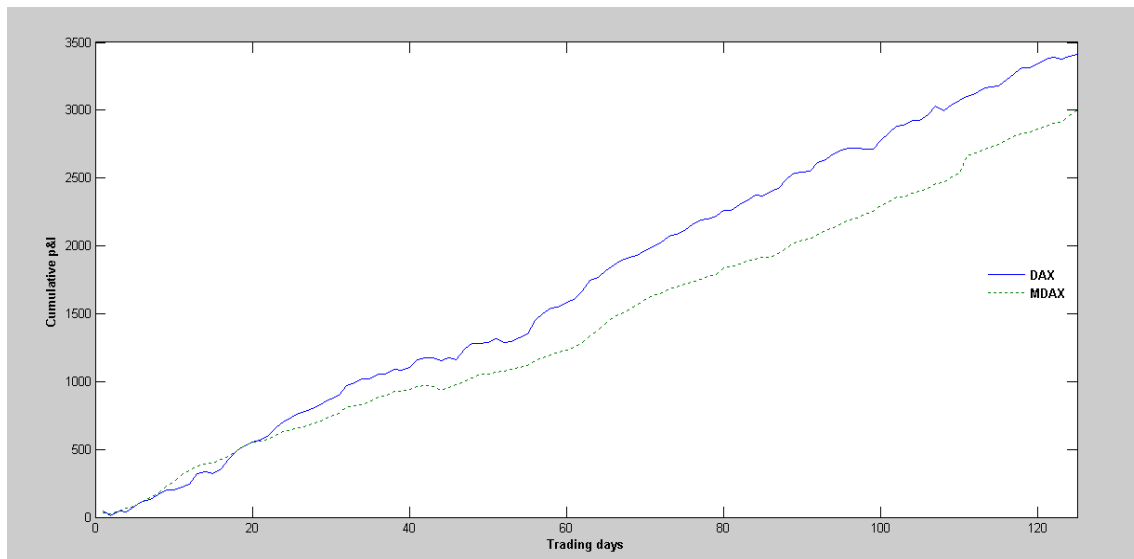
Horizontal axis: data sample of 125 trading days; vertical axis: number of occurrences of UEEs per day, DAX events are positive and MDAX events are negative for presentation purposes.

Figure 1. 3 Number of UEEs per minute: DAX - MDAX



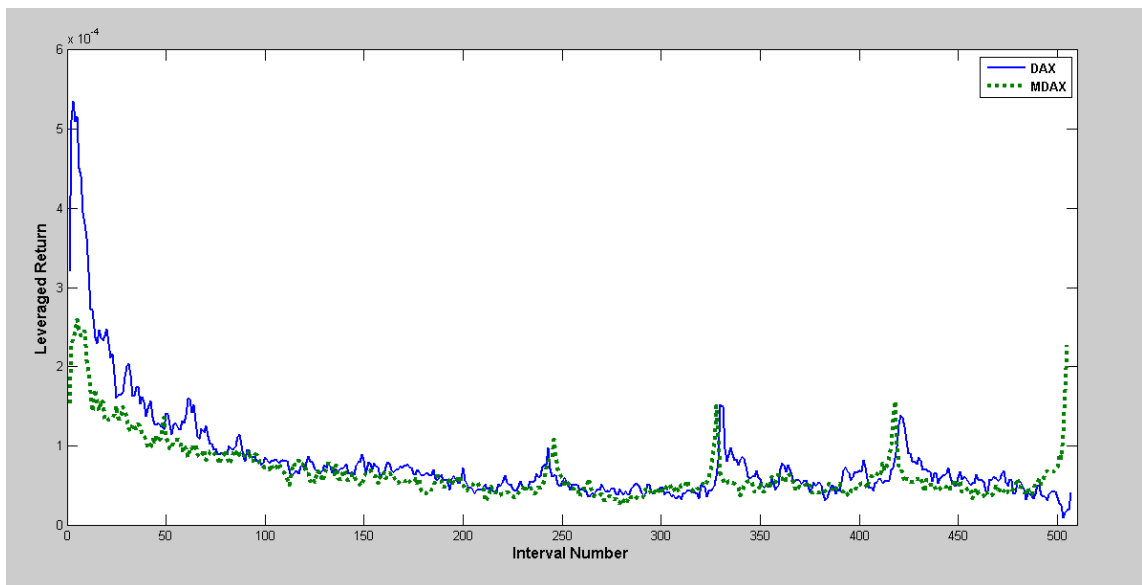
Horizontal axis: data sample of 510 minutes of trading per day; vertical axis: number of occurrences of UEEs per minute, DAX events are positive and MDAX events are negative for presentation purposes.

Figure 1. 4 DAX - MDAX cumulative P&L: OQP



Aggregated cumulative Profits & Losses obtained by Ait-Sahalia and Saglam (2014) Optimal Quoting Policy over the sample period (February 2 to July 30, 2013).

Figure 1. 5 DAX- MDAX leveraged return per time interval



One-minute time-volume weighted average returns (TVWAR) obtained by the trading strategy. Returns are aggregated by market indexes.

Chapter 2

Identification of algorithmic imprints using public data: an application to batch auctions

Abstract

We identify and classify algorithmic trading activities with batch auction data from Xetra, the electronic platform of the German stock exchange. We pinpoint negative price loops' high-frequency activities by combining Yunyue and Shasha (2003) shifted wavelet tree (SWT), a burst detection indicator, and the dynamic time warping (DTW) similarity measure of Skutkova, Vitek et al. (2013). Then, with Kirilenko, Kyle et al. (2016) traders classification, we link algorithmic activities to their behavioral characteristics.

Nine types of algorithm imprints are identified. The first five types are negative price loops. They exhibit the encrypted noise characteristics (Stiglitz 2014) and they hamper the price discovery process. They significantly modify the liquidity information sets. The last four types are behavior-based. Informed traders imprints are documented from both high and low-frequency traders. The economic value of algorithm detection revolves around two axes. Market participants can adapt their trading to the presence of encrypted noise by filtering data and clarifying the price discovery process. Deciphering of NPLs encrypted noise reveals algorithmic imprints from the informed traders. This can lead to significant economic gains. The methodology deployed makes it possible to adapt to different environments, including continuous trading.

Introduction

Batch auctions (the auctions hereafter) are widely used at the opening and closing of stock markets. Adopted by almost all stock exchanges around the world, they are crucial trading mechanisms. The beneficial effect of auctions on market quality is well documented by means of event studies associated with the inauguration of auctions on various stock exchanges. The Paris stock exchange implemented them on two segments successively, one in 1996 and the other in 1998. Pagano and Schwartz (2003) determine that auctions lower execution costs for participants. The Singapore Exchange enforced opening and

closing auctions in August 2000. Comerton-Forde, Ting Lau et al. (2007) and Chang, Rhee et al. (2008) conclude to an improvement in market quality and a decrease in end-of-day price manipulation. Closing auction started at the London Stock Exchange in May 2000. Chelley-Steeley (2008) notices an improvement in market quality at the exchange. Pagano, Peng et al. (2013) analyze the impact on bid-ask spread and price volatility of auctions introduced in 2004 on NASDAQ. Their results suggest positive spillovers on the price formation dynamic behavior. In June 2009, Nasdaq-OMX launches index futures auctions. Hagströmer and Nordén (2014) conclude that they improve closing price accuracy and end-of-day liquidity.

Algorithmic trading (AT), of which high-frequency trading is a subset, is defined as a trading system whose decision-making process does not involve human intervention (Bates (2017)). It is the expression of a fundamental trend centered on technological development. The nature of competition evolves as high-frequency traders change speed into information (O'Hara (2015)). J.P. Morgan (2017) notes that for short-term trading, humans already play a very small role. During continuous time sessions, Bouveret, Guillaumie et al. (2014) estimate the value traded by high-frequency traders to 24% in Europe and 21% on Xetra. AT accounts for approximately 50% of liquidity demand and supply on the Deutsche Boerse (Hendershott and Riordan (2011)). Despite its importance, research papers focusing on AT during continuous trading are scarce. Menkveld (2016) identifies two papers using public data. Hendershott, Jones et al. (2011) define the message rate as a proxy of AT activities and do not differentiate between high and low-frequency traders. Latza, Marsh et al. (2014) subdivide trades according to their reaction times. They classify an aggressive order as “fast” if it matches against a standing limit order that is less than 50 milliseconds old. It is “slow” otherwise. They conclude that fast trades have smaller execution costs than slow trades. To the best of our knowledge, no study documents the AT presence and its behavior during auctions.

We identify and classify trading algorithm activities from the Frankfurt stock exchange auction data. Information opacity is more important than that of continuous time. We present a methodology to infer the orders characteristics. Our algorithmic imprints recognition exploits two approaches. First, we focus on Abrol, Chesir et al. (2016)

negative price loops (NPLs). Public information flow is monitored to detect high-frequency activity bursts. Yunyue and Shasha (2003) shifted wavelet tree (SWT) structures the data and we apply a burst indicator to reveal activity eruptions. Skutkova, Vitek et al. (2013) dynamic time warping (DTW) compares the burst sequence to a pre-defined NPL sequence to determine their similarity. We identify five types of NPLs: 456,772 events are uncovered representing more than 11% of all auction events. The NPLs users can either be informed traders or proprietary firms. NPLs exhibit the Stiglitz (2014) encrypted noise characteristics and they blur the price discovery process. NPLs significantly influence the liquidity information sets. Second, we classify algorithm types according to Kirilenko, Kyle et al. (2016) behavioral characteristics. Four algorithm types are identified. They differ by their impact on the price discovery process, the aggressiveness of their orders, and their position management. These behavioral sequences account for 596,220 events and represent more than 14% of all activities. A strong concern to minimize the impact on indicative price, the type of position management, and the temporal cyclicity documented denote the presence of informed traders using high and low-frequency infrastructures.

The paper structure is as follows: Section 2.1 presents a review of the literature. Section 2.2 characterizes the institutional context of trading on Xetra, the electronic platform of the German stock exchange. Section 2.3 defines the concept of algorithmic sequences and develops the methodologies allowing their identification. Section 2.4 presents the data and transaction costs. Section 2.5 and Section 2.6 outline and discuss the results. Section 2.7 concludes.

2.1 Literature review

2.1.1 Behavior

Numerous studies investigate the behavior of high-frequency traders (HFTers) during continuous double auctions. Brogaard, Carrion et al. (2016) conclude that HFTers supply liquidity during extreme price movements. Subrahmanyam and Zheng (2016) note the ability of HFTers to manage limit orders in anticipation of short-term price movements. Goldstein, Kwan et al. (2016) find that HFTers provide liquidity on the thick side of the

order book and demand liquidity on the thin side. Hirschey (2016) states that HFTers can anticipate the order flow from other investors. Menkveld and Yueshen (2016) emphasize the importance of inter-market arbitrage as a behavioral characteristic. These studies allow a better understanding of the industry aggregate behavior but do not distinguish between specific traders activities who may exhibit heterogeneous behaviors (Carrion (2013))³. There are exceptions: Menkveld (2013) highlights the positive contribution of a high-frequency market maker (HFMM) arrival on Chi-X Europe, and Yergeau (2016) analyzes the behavior of an endogenous liquidity provider using the dynamic inventory management model of Ait-Sahalia and Saglam (2014).

Kirilenko, Kyle et al. (2016) propose a trader behavioral classification based on their activities and inventory management type. The authors benefit from a complete observation of all market participants` activities from a Commodity Futures Trading Commission audit database. Table 2.1 describes the types of active traders on the E-mini S&P 500 futures contract of the Chicago Mercantile Exchange.

[insert Table 2.1 here]

They observe that intraday intermediaries share common characteristics. Positions are small relative to the limit order volumes and inventories exhibit mean reversion. The authors distinguish two types of traders in this category. The HFMMs have inventories negatively correlated to stock prices while inventories are positively correlated to stock prices for HFTers. Fundamental traders trade large quantities of which a minimum of 15% remains in inventory. Their positions are directional.

2.1.2 Machine learning

Yang et al. (2012) utilize machine learning to identify the Kirilenko, Kyle et al. (2016) categories of traders. Variables at the origin of all decisions are the inventory position (Kyle (1985); Glosten and Milgrom (1985); Huang and Stoll (1997); among others) and

³ A review of the high-frequency trading industry: Chung, K. H. and A. J. Lee (2016). "High-frequency Trading: Review of the Literature and Regulatory Initiatives around the World." Asia-Pacific Journal of Financial Studies **45**(1): 7-33.

the imbalances at the first and third levels of the order book (Cont, Kukanov et al. (2014)). The authors wish to obtain trader categories reward functions from inverse reinforcement learning. Eighteen simulations of approximately 300,000 E-Mini S&P 500 LOB activities serve as learning. Simulations come from Hayes, Paddrik et al. (2012) agent-based model. They show a clear connection between the traders classification from Kirilenko, Kyle et al. (2016) and the results from their machine learning approach.

2.2 Institutional context: trading on Xetra

Table 2.2 shows that a hybrid market model with three auctions and two continuous trading periods are in effect for the DAX and MDAX segments. During auctions, traders can submit limit and market orders. At auction's end, matching orders are executed at a single price and unexecuted orders transferred to the next trading stage.

[insert Table 2.2 here]

[insert Table 2.3 here]

Table 2.3 shows the available public information during auctions. Public information consists of eight elements: the stock identifier, the date, the timestamp in microseconds, the status of the auction (opening, intraday or closing), the indicative price, the quantity matched at the indicative price, the side of the surplus (imbalance) and its quantity. The identity of the trader, the type of event (creation, modification, or cancellation), and the quantity associated with each event are not published. They are deduced from public information with interpretations in Table 2.4.

[insert Table 2.4 here]

Any matched quantity variation represents the order's quantity. If the matched quantity increases and the indicative price increases (decreases), it is a buy (sell) order. If the matched quantity decreases and the indicative price increases (decreases), this is a cancellation of a sell (buy) order. In absence of a matched quantity change, the surplus variation corresponds to the event quantity. An increase in the surplus variation is due to

a limit order creation on the side of the marginal variation whereas a decrease is due to a cancellation.

Since our data does not identify orders specifically, our methodology no doubt introduces noise in the association of activity sequences to a given source. This bias could originate from the aggregation of two or more orders reported during the same event. Even if possible, microsecond timestamps suggest that our methodology can infer related algorithmic sequences.

2.3 Algorithmic sequences

Our goal is to link algorithmic sequences to algorithm types. Hasbrouck and Saar (2013) correlate same quantity limits and/or market orders with short duration to high-frequency algorithms. We apply this concept to auctions. First, we identify algorithmic sequences with the SWT tree structure (Yunyue and Shasha (2003)) and a burst detection indicator. DTW (Skutkova, Vitek et al. (2013)) measures the similarity of these sequences to reference sequences exhibiting NPLs' characteristics documented in the literature. Second, we classify the algorithmic sequences using the behavioral characteristics of Kirilenko, Kyle et al. (2016). We focus on two attributes. Orders' aggressiveness determined by their impact on the indicative price and orders' cancellation rates which reveal real intention to trade.

2.3.1 Negative price loops (NPLs)

Yunyue and Shasha (2003) propose a tree structure, the SWT, to monitor bursts of real-time activities from a data stream. The tree aggregates the time intervals while preserving the original structure of the data. The main contribution of Yunyue and Shasha (2003) is the reduction in the number of windows necessary to monitor events. Their structure shrinks from $O(n^2)$ by considering the set of all possible combinations to $O(n)$ where n is the number of windows of the smallest time interval of the sample considered. Figure 2.1 illustrates the structure of the SWT which exploits the half-overlap of time windows.

[insert Figure 2.1 here]

Equation (2.1) uses the duration and number of SWT levels to obtain the basic time interval of the tree, i.e. the level 0 time interval:

$$SWTtime\ interval_{level0} = \frac{time_{end} - time_{start}}{2^l}, \quad (2.1)$$

where: l = number of SWT levels; $l \in [1..L]$; $L \in \mathbb{Z}^+$.

For a closing auction with a total duration of five minutes, we use a 14-level tree with 16,384 windows (2^{14}). The time interval at level 0 is 0.0183 second ($(5\ minutes * 60)/16,384$). The tree reduces the number of windows to supervise from 268,435,456 ($16,384^2$) to 16,384 windows. Burst detection of abnormal activities in a timely manner becomes feasible. Their identification depends on the intraperiod cumulative value of the burst indicator $F(\cdot)$ and a threshold:

$$F(x_{i,j}) > f(w_i), \quad (2.2)$$

where:

$x_{i,j}$ = time interval of SWT level i , window j ,

$f(w_i)$ = time interval threshold of SWT level i .

$f(w_i)$, the alarm's domain, is equal to $\min(6, 2^i)$, i being the SWT level monitored.

When an alarm comes from a higher level, efficient streaming algorithms (online and batch) are available from Yunyue and Shasha (2003) *Lemma 3*. This makes it possible to precisely locate the level 0 sequence involved.

In order to link the events of activity bursts to a specific type of algorithm, we compare them to reference sequences identified using stylized facts (quote stuffing: Egginton, Van Ness et al. (2014); Ye, Yao et al. (2013); Cr dit Suisse (2012); Brogaard (2010), and phantom liquidity: Blocher, Cooper et al. (2016); Korajczyk and Murphy (2016)).

We quantify the similarity between reference sequences and sequences from alarms with the DTW's distance. DTW finds an optimal alignment between two data sequences. It

minimizes time shift and distortion effects. It can measure the similarity between two series that may differ in length. This similarity measure is the best in pattern recognition (Petitjean, Forestier et al. (2014); Ding, Trajcevski et al. (2008)). We perform data mapping $\in \{-1,1\}$ so that the temporal similarity has a meaning (Keogh and Kasetty (2003)).

2.3.2 Behavior-based strategies

We use the percentage of activities at the indicative price and the orders cancellation rate to define four categories presented in Table 2.5. The categories qualify behavior (aggressive or conservative), impact on price discovery process (positive or neutral), and type of position management (directional or not) that the algorithmic sequences exhibit.

[insert Table 2.5 here]

During auction, two types of activity affect public information: bid and ask orders (created or canceled) at the indicative price (referred to as *locks*) and bid (ask) orders created or canceled whose price is higher (lower) than the indicative price (referred to as *crosses*). Moshirian, Nguyen et al. (2012) link the aggressiveness of an order with the probability that the order submitted is executed if the market opened at the event arrival time. In this sense, all activities generating public information during the auction are aggressive. However, Cao, Ghysels et al. (2000), Rinaldo (2004) and Anagnostidis, Kanas et al. (2015) precise the nature of these activities: *crosses*, by their impact on prices, imply more aggressive signals than *locks*. We define an algorithmic sequence as aggressive if fewer than 70% of its events occur at the indicative price. Otherwise, the sequence is conservative.

The orders cancellation rate identifies activities issued by high-frequency traders (Paddrik, Haynes et al. (2016); O'Hara (2015); Aitken, Cumming et al. (2015)). A low cancellation rate indicates real intention to execute orders, a characteristic attributed to informed traders in the sense of Glosten and Milgrom (1985)⁴. Behavioral-based

⁴ “The informed trader may be speculating based on private information or superior analysis, or he may simply have a ‘liquidity’ reason for trading, but in any event, his decision to buy, sell or leave is based on his information.” p. 77.

algorithmic sequences (behavioral sequences hereafter) include high and low-frequency activities. They are limit and/or market orders of similar quantities which meet the following criteria:

- 1) A minimum number of 4 events per auction. Institutional investors⁵ use algorithms to implement their strategies (BlackRock (2014)). They are considered as informed traders in the literature (Choi, Jin et al. (2016)). The low number of repetitions is intended to preserve the ability to identify this type of traders.
- 2) Sequence median duration is less than one second.
- 3) Maximum duration is shorter than five seconds. An event at time t having duration greater than the maximum duration followed at $t + 1$ by an event whose duration is less than the maximum duration causes a new sequence creation.

2.4 Data

Data comes from Xetra, the electronic platform of the Frankfurt Stock Exchange. The database contains all events relative to auctions sent via the Enhanced Broadcast System, an information flow used by high-frequency traders. Xetra Parser, developed by Bilodeau (2013), allows to reconstruct the events sequence. The timestamps are in microseconds and trading is anonymous.

Our sample has sixty components: thirty from the DAX index, the stocks with the highest market capitalization and thirty from the MDAX index, the stocks having an average market capitalization and excluding technology. Hereinafter, we refer to the thirty DAX (MDAX) components as DAX (MDAX). The sample covers the period from February to July 2013. It accounts for about 15% of all activities, i.e. 4 094 751 events. The other 85% essentially occurs during continuous trading sessions.

Table 2.6 shows the statistics by auction.

⁵ Brogaard, J., et al. (2014). "High-Frequency Trading and the Execution Costs of Institutional Investors." *Financial Review* **49**(2): 345-369.

: "Institutional investors refer to buy-side institutions such as pension plans and money managers."

[insert Table 2.6 here]

Closing auctions trigger most activities with 71% (69%) of the DAX (MDAX) events. The relative importance of the three auctions are qualitatively the same in both indexes. The bulk of activities happen during closing auctions. Differences between average and median matched quantities is due to the presence of frequent extreme values.

Table 2.7 classifies events according to their impact on matched quantity. 15.03% (35.52%) of DAX (MDAX) events increase the matched quantity. This is the result of limit orders decreasing the existing surplus size, a conservative strategy, or aggressive limit orders affecting the indicative price. Quantity additions to existing surplus represent the majority of events: 52.78% (DAX) and 51.05% (MDAX). This reflects a concern to minimize orders impact on indicative price, an institutional traders characteristic. Cancellation of previously matched orders accounts for 1 062 314 (106 576) DAX (MDAX)'s events.

[insert Table 2.7 here]

Easley, Lopez de Prado et al. (2012) link temporal cyclicity to institutional traders. Figure 2.2 shows the behavior of DAX closing auctions event numbers by 5-second time intervals. Six bursts in the number of orders occur at the same time for all components of the Index⁶. This is a clear evidence of institutional imprints. The very low activity observed during the auctions' last thirty seconds (periods 61 to 66) meaning that random time period addition at auctions' end has, at best, a mixed economic contribution. These behaviors are also seen for other auctions and the MDAX.

[insert Figure 2.2 here]

2.4.1 Transaction costs

There are many types of traders. Hedgers and institutional investors have heterogeneous investment horizons (Cespay and Vives (2016)). Some may use brokers for their orders execution. If so, Battalio, Corwin et al. (2016) identify US brokers who maximize their

⁶ We obtain the same cyclicity pattern when we split the sample in shorter subsamples.

revenues by the rebates granted by trading venues to the detriment of their customers. Deutsche Boerse (2015) states that under the Designated Sponsor Program (Section 2.2.3.2) and the Top Liquidity Provider Program (Section 2.2.3.3), the stock exchange does not charge transaction fees to participants to these programs and grants them rebates for executed orders (limit and marketable). Top liquidity providers earn a rebate of 0.20 basis point on their traded market value cap to 375k euros per order per day. Other hedgers and institutional investors can benefit from direct market access (DMA). Some of them may use co-location (Malinova and Park (2016); Malinova, Park et al. (2016)). Direct access allows the management of limit and marketable orders strategically using several prices (Upson and Van Ness (2017); Easley, de Prado et al. (2016)). If these investors are billed directly by the Deutsche Boerse, the Section 2.2.1.1, Table 6 (Deutsche Boerse (2015) *op.cit.*) establishes for the DAX a fee model based on three activity levels referred to as high volume, medium volume and low volume levels. The cost of the medium volume category is 0.378 basis points based on the market values capped to 1.5m euros per order per day. This model does not qualify for rebates. In the case of billing by the broker, Cappon and Mignot (2014) estimate costs at 1.5 cents per share, while Menkveld (2016) estimates the cost of executing a marketable order at 7 basis points.

The data during auction is opaque. The orders issuers and their characteristics are not in the public domain. We do not identify the exact status of the traders behind the algorithmic activities.

2.5 Results

2.5.1 Negative price loops

NPLs lock the indicative price in a range. They prevent the price discovery process and hamper the disclosure of supply and demand (Abrol, Chesir et al. (2016)). NPLs may create phantom liquidity, a source of additional costs imposed on investors (BMO Capital Markets (2009)). They are associated to quote stuffing, a cause of latency's arbitrage (Brogaard (2010)) and stale quotes (Foucault, Kozhan et al. (2015); Menkveld and Zoican (2017)).

SWT and DTW identify five types of NPL algorithms whose characteristics are presented in Table 2.8. All NPL algorithms share a common structure. An uninterrupted creation and cancellation sequence of identical orders (price and quantity) implying cancellation rates which converge to 100%. They use limit or market orders. They set the indicative price in a constant range. We categorize NPLs algorithms by their impact on information. The impact depends on heterogeneous (LOB's depth and granularity) and endogenous (order's price and quantity) factors. Types 1 and 2 have the smallest effect. Limit orders are inside the bid-ask spread on the surplus side and they don't modify the matched quantity, i.e. they don't match with the LOB's opposite side. Types 3 and 4 have a greater impact as they change the matched quantities. The imbalance (surplus quantity) stays on the same side of the market. This can be the result of a marketable order with a quantity smaller than the surplus. Type 5 influences all variables.

[insert Table 2.8 here]

NPLs execution speed does not allow humans to perceive their activities. A graphic presentation whose paradigm is event-driven rather than temporal is revealing. Figure 2.3 displays Deutsche Bank closing auction events on 12 February 2013. The graphs show clockwise: the indicative price, the matched quantity, the surplus quantity, and the duration between events. We identify four NPL sequences. As previously defined, the indicative price is within a constant range. Indicative price volatility changes with sequences. Duration is significantly shorter during these NPLs than on average and confirm activities related to high-frequency algorithms.

[insert Figure 2.3 here]

Figure 2.4 illustrates the analyzed variables for NPLs identification. An identical order (price and quantity) is created and canceled six times (twelve events) during a time interval of 0.08 seconds without any other activity intervening. We link the sequence to Table 8 type 1 algorithm: indicative price and sell side surplus vary while the matched quantity remains the same. Duration between events range from 0.4 to 2.0 milliseconds. To act at this speed, the algorithm is probably operated from a collocation site. DTW determines the similarity between reference sequences and the potential NPL sequence.

[insert figure 2.4 here]

403,746 DAX events are due to NPLs (Table 2.9). These mainly arise at the closing auction where they represent more than 17% of events. These high-frequency events occur at the first SWT levels: 61.71% of the activities are detected at level 2 where the time interval is 0.08 seconds. Each increase in one SWT level doubles the preceding time interval. The MDAX behavior is qualitatively similar.

[insert Table 2.9 here]

[insert Table 2.10 here]

Table 2.10 summarizes the NPLs characteristics. We identify 2,595 DAX NPLs sequences. They average 143 repetitions and have a median duration of 0.012 seconds. Sequences are mainly arising during closing auction. MDAX has 134 NPL sequences affecting eight index components. These sequences occur at the end of the day and exhibit 357 repetitions on average.

[insert Figure 2.5 here]

Figure 2.5 shows the NPLs distributions during the DAX and MDAX closing auctions. NPLs are present from the second to the penultimate minute of these auctions. Figure 2.6 displays the distribution of NPLs events for the five algorithm types of Table 2.8. They differ significantly. For the DAX, all types are used and almost one third of the NPLs (129,823 events) are generated by algorithm 5. This algorithm is by far the most aggressive by its characteristics. It implies a change in the surplus side as well as variations in the indicative price and the matched quantity. For the MDAX, only types 1 and 2 are used. These types place limit orders inside the bid-ask spread.

[insert Figure 2.6 here]

2.5.2 NPLs and liquidity

To determine whether the NPLs significantly influence the liquidity information set, we examine their effects on the liquidity distributions including and excluding the NPLs. We

define the available liquidity at event i as the positive (negative) value of the imbalance if the surplus is on the buyer (seller) side.

The stable distributions family adequately defines asymmetric and leptokurtic distributions using parameters λ for stability, β for skewness, γ for scale, and δ for location (Vidyasagar (2014); Barany, Varela et al. (2012); Sewell (2011)). As no closed form exists for the pdfs and the cdfs except for the Gaussian ($\lambda = 2$), the Cauchy ($\lambda = 1$, $\beta = 0$) and the Levy ($\lambda = 0.5$, $\beta = 1$) distributions (Cizek, Härdle et al. (2011)), we fit the four parameters to the empirical characteristic function using the regression-type method from Koutrouvelis (1980). Matlab programming is from Li (2015).

Table 2.11 contains the estimated parameters of the stable family distributions for the DAX components. α and β determine the shape of the distributions (Nolan (2015)). The presence of NPLs has the following implications on the shape of liquidity distributions: α variations are in the range 2.31% to 38.39%. The liquidity distributions without NPLs have lower peak and lower tails. β variations fluctuate widely with values between -347.29% and 3,328.68%. HFTers implement a large span of strategies. The aggressiveness and the LOB side of the strategy dictate the NPLs effect on skewness.

[insert Table 2.11 here]

Table 2.12 shows the results from a Kolmogorov-Smirnov two samples nonparametric test. The similarity between the empirical cumulative distributions with and without NPLs is rejected at the 1% confidence level for all stocks. NPLs significantly modify the liquidity information set. NPLs are algorithmic sequences of creations and cancellations of similar orders whose durations are ultra-short. Although the initial order creation and cancellation sequence can provide information about the hidden order book depth and its granularity, the subsequent repetitions are redundant.

[insert Table 2.12 here]

Similar results are obtained for the MDAX components.

2.5.3 Behavioral-based strategies

Table 2.13 presents the behavior-based sequences statistics. They represent a minimum of 12.45% of the DAX events. For the MDAX, behavioral sequences account for 23.15% of closing auction events. Overall, NPLs depict 596,220 events.

[insert Table 2.13 here]

Table 2.14 groups the behavioral sequences by market segments and by algorithm types. For the DAX, 412,782 events, or more than 90% of the 457,864 events associated with behavioral sequences, occur at the indicative price (types 6 and 8). Of these, 119,740 orders have a cancellation rate smaller than 70% (type 8). This suggests the presence of informed traders managing their positions with algorithms. This characteristic is also true for the MDAX where 91.4% of behavioral sequences are of type 8.

[insert Table 2.14 here]

Table 2.15 and Table 2.16 identify the characteristic that predominantly differentiates the behavioral types. The percentage of limit orders at the indicative price (types 6 and 7) is greater than 90% for all DAX and MDAX auctions. Operators of these algorithm types exhibit a strong concern to minimize their orders' impact on the stock fundamental value.

[insert Table 2.15 here]

Position management imprints (types 8 and 9), also associated with informed traders, are present during all DAX auctions. 132,852 events are identified. Here again, the highly-privileged strategy is to minimize the impact on the indicative price (type 8): it represents more than 90% of total position management activities. For the MDAX, nearly 97% of the 138,356 events identified are of Type 8. They occur during the closing auction.

[insert Table 2.16 here]

2.6 Discussion

Orders anticipation is an integral part of the strategies commonly used by high-frequency traders. Baldauf and Mollner (2016) show that high-frequency liquidity providers use

order identification to avoid adverse selection by canceling mispriced quotes. Brunnermeier and Pedersen (2005) describe as predatory the exploitation of orders from institutional investors having to liquidate their positions. Clark-Joseph (2013) documents the behavior of HFTers that probe liquidity in order to obtain information about large incoming orders. Yang and Zhu (2016) introduce the concept of back-running in which fast traders compete with the institutional investors large orders after recognizing their imprints. Thus, each participant has an economic incentive to make it more difficult for other traders to extract information from public data (the Stiglitz (2014) data encryption hypothesis).

SWT and DTW allow to reveal NPLs. They abound during auctions. NPLs are a by-product of low-latency trading as they are characterized by short durations. Position management is not their goal, their cancellation rates converge to 100%. They are conceptualized to generate redundant information that does not improve market quality. They blur the price discovery process and significantly modify the liquidity information set. NPLs are consistent with Stiglitz (2014) data encryption hypothesis. Institutional investors and proprietary firms use AT (Hasbrouck and Saar (2013); Yang and Zhu (2016)). As we cannot identify the order's originator, we must consider that the presence of NPLs is due to either of two scenarios. First, the informed trader wish to conceal his trading intention. Second, the proprietary firms induce delays in information processing justified by their desire to take advantage of information extracted from the informed traders. Figure 2.7 shows a histogram of the NPL sequences transaction costs. Transaction cost per sequence is 56.70 euros maximum⁷ by NPL sequence. These costs do not deter traders to encipher the price discovery process. Moreover, if the NPLs' generators are part of the designated sponsor or top liquidity provider programs (Deutsche Boerse (2015), Section2.4.1), they do not incur transaction costs. They earn rebates.

[insert Figure 2.7 here]

⁷ 1.5m euros multiplied by 0.378 basis point as we refer to medium volume activities fees from Deutsche Boerse (2015). "Price List for the Utilization of the Exchange EDP of FWB Frankfurt Stock Exchange and of the EDP XONTRO."

The regulators face numerous challenges with encrypted noise. The stock exchange encourages their use by high-frequency traders with its fee and rebate structure. Encrypted noise ban may hamper low-latency informed traders to execute their trades with a minimum of price impact. Designing regulations to forbid encrypted noise by constraining orders submission can be difficult because it opens up opportunities for regulatory arbitrage (Stiglitz (2014)). Tackling the problem by reducing trading speeds would also affect liquidity takers (Shorter and Miller (2014)). If regulatory bodies target specific behaviors, quants can modify the algorithms.

NPLs have cancellation rates converging to 100%. For both DAX and MDAX, the main concern of the identified sequences is to limit the orders impact on the indicative price, a characteristic of large investors during continuous time trading (Duffie and Zhu (2017)). Figure 2.8 embodies this preoccupation. A time weighted average price (TWAP) algorithm is active during the 15 May 2015 closing auction on Volkswagen stock (identifier 130). Seven events occur in a 0.073 second timeframe. We comment the graphics clockwise. None of the seven trades modifies the indicative price. No surplus on the offer side. The short position increases steadily by a quantity of 20 to end at -140. It is the result of seven marketable orders sent against the bid surplus quantity. All duration is in the 0.011 - 0,015 second range. Such regularities require a low latency infrastructure. This behavior is in line with Menkveld (2016) who documents patterns of liquidity-demanding tradebots from high and low-frequency traders.

[insert Figure 2.8 here]

The criteria for identifying behavioral algorithms are based upon the aggressive orders use of lock and cross types (Section 2.3.2). The analysis of behavioral algorithmic sequences suggests that dynamic order management is identical to that used in continuous time. The informed traders are concerned about the impact of their orders on the indicative price and they resort to fragmentation. For example, a repositioning in the order book on the side of the surplus is done by mean of cancellation and creation of new orders of types lock or cross. Another behavior implies the use of crosses orders consuming part of the

surplus followed by locks or crosses on the same side of the surplus for identical or greater quantities following indicative price variations.

The opacity of the available information prevents the tractability of the orders which would allow to quantify the economic impact obtained by the algorithms. However, the economic value of algorithm detection revolves around two axes. First, market participants can adapt their trading to the presence of encrypted noise. By defining appropriate burst indicators and targeting specific timeframes, SWT and DTW can accustom themselves to monitor encrypted noise evolution in streaming and batch environments. Filters can be implemented to mitigate data congestion while clarifying the price discovery process. Second, deciphering of NPLs encrypted noise reveals algorithmic imprints from the informed traders. This can lead to significant economic gains.

2.7 Conclusion

In this study, we present the first direct evidence of algorithmic imprints during auctions. Our sequential approach makes it possible to recognize algorithm types. They account for more than one million events (a quarter of the total). While NPL algorithms informational content varies, they all hamper the information processing speed and significantly modify the liquidity information set. The identification of NPLs facilitates the isolation of behavior-based sequences exhibiting real intent to trade. A strong concern to minimize the impact of orders on the indicative price, position management features and the sudden and periodic burst of activities suggest the presence of informed traders who leave identifiable imprints. Comparing the two indexes, position management differs. A greater trading intensity coupled with hedging operations and arbitrage opportunities between the DAX components and the very liquid DAX futures may be at the origin of this difference.

The economic value of algorithm detection revolves around two axes. Market participants can adapt their trading to the presence of encrypted noise by filtering data and clarifying the price discovery process. Deciphering of NPLs encrypted noise reveals algorithmic imprints from the informed traders. This can lead to significant economic gains. The methodology deployed makes it possible to adapt to different environments, including continuous time trading.

References

- Abrol, S., et al. (2016). "High Frequency Trading and US Stock Market Microstructure: A Study of Interactions between Complexities, Risks and Strategies Residing in U.S. Equity Market Microstructure." Financial Markets, Institutions & Instruments **25**(2): 107-165.
- Ait-Sahalia, Y. and M. Saglam (2014). "High Frequency Traders : Taking Advantage of Speed." SSRN Electronic Journal.
- Ait-Sahalia, Y. and M. Sağlam (2016). "High Frequency Market Making." SSRN Electronic Journal.
- Aitken, M., et al. (2015). "High Frequency Trading and End-of-day Price Dislocation." Journal of Banking & Finance **59**: 330-349.
- Anagnostidis, P., et al. (2015). "Information Revelation in the Greek Exchange Opening Call: Daily and Intraday evidence." Journal of International Financial Markets, Institutions and Money **38**: 167-184.
- Baldauf, M. and J. Mollner (2016). "Fast Traders Make a Quick Buck: The Role of Speed in Liquidity Provision." SSRN Electronic Journal.
- Barany, E., et al. (2012). "Detecting Market Crashes by Analyzing Long-memory Effects Using High-frequency Data." Quantitative Finance **12**(4): 623-634.
- Bates, T. (2017). "Topics in Stochastic Control with Applications to Algorithmic Trading." LSE - Ph. D. Thesis.
- Battalio, R. H., et al. (2016). "Can Brokers Have It All? On the Relation between Make-Take Fees and Limit Order Execution Quality." Journal of Finance, Forthcoming.
- Bilodeau, Y. (2013). "Xetraparser [computer software]."
- BlackRock (2014). US Equity Market Structure: An Investor Perspective: 1-8. Research Report.
- Blocher, J., et al. (2016). "Phantom Liquidity and High Frequency Quoting." Journal of Trading **11**(3): 6-15.
- BMO Capital Markets, Q. E. S. (2009). "The Hidden Costs of Locked Markets." Research Report.
- Bouveret, A., et al. (2014). "High-frequency Trading Activity in EU Equity markets." ESMA Economic Report.
- Brogaard, J. (2010). "High Frequency Trading and its Impact on Market Quality." SSRN Electronic Journal.
- Brogaard, J., et al. (2016). "High-Frequency Trading and Extreme Price Movements." SSRN Electronic Journal.

- Brogaard, J., et al. (2014). "High-Frequency Trading and the Execution Costs of Institutional Investors." Financial Review **49**(2): 345-369.
- Brunnermeier, M. and L. H. Pedersen (2005). "Predatory Trading." Journal of Finance **60**(4): 1825-1863.
- Cao, C., et al. (2000). "Price Discovery without Trading: Evidence from the Nasdaq Preopening." The Journal of Finance **55**(3): 1339-1365.
- Cappon, A. and S. Mignot (2014). "The Brokerage World Is Changing, Who Will Survive?" Forbes **April 16**(<https://www.forbes.com/sites/advisor/2014/04/16/the-brokerage-world-is-changing-who-will-survive/#68e93b9068a7>).
- Carrion, A. (2013). "Very Fast Money: High-frequency Trading on the NASDAQ." Journal of Financial Markets **16**(4): 680-711.
- Cespay, G. and X. Vives (2016). "The Welfare Impact of High Frequency Trading." American Finance Association - Conference.
- Chang, R. P., et al. (2008). "How Does the Call Market Method Affect Price Efficiency? Evidence from the Singapore Stock Market." Journal of Banking & Finance **32**(10): 2205-2219.
- Chelley-Steeley, P. L. (2008). "Market Quality Changes in the London Stock Market." Journal of Banking & Finance **32**(10): 2248-2253.
- Choi, J. J., et al. (2016). "Informed Trading and Expected Returns." Yale University, Working Paper.
- Chung, K. H. and A. J. Lee (2016). "High-frequency Trading: Review of the Literature and Regulatory Initiatives around the World." Asia-Pacific Journal of Financial Studies **45**(1): 7-33.
- Cizek, P., et al. (2011). "Statistical Tools for Finance and Insurance". Elsevier Editor. 410 pages.
- Clark-Joseph, A. D. (2013). Exploratory Trading. Harvard Working Paper.
- Comerton-Forde, C., et al. (2007). "Opening and Closing Behavior Following the Introduction of Call Auctions in Singapore." Pacific-Basin Finance Journal **15**(1): 18-35.
- Cont, R., et al. (2014). "The Price Impact of Order Book Events." Journal of Financial Econometrics **12**(1): 47-88.
- Crédit Suisse (2012). What Does "Bad" HFT Look Like, How often Does It Happen, and How Do We Detect It? AES Analysis.
- Deutsche Boerse (2015). "Price List for the Utilization of the Exchange EDP of FWB Frankfurt Stock Exchange and of the EDP XONTRO." Research Paper.
- Ding, H., et al. (2008). "Querying and Mining of Time Series Data: Experimental Comparison of Representations and Distance Measures." PVLDB **1**(2): 1542-1552.

- Duffie, D. and H. Zhu (2017). "Size Discovery." The Review of Financial Studies **30**(4): 1095-1150.
- Easley, D., et al. (2016). "Discerning Information From Trade Data." Journal of Financial Economics **120**(2): 269-285.
- Easley, D., et al. (2012). "The Volume Clock: Insights into the High-frequency Paradigm." Journal of Portfolio Management **39**(1): 19-29.
- Egginton, J. F., et al. (2014). "Quote Stuffing." SSRN Electronic Journal.
- Foucault, T., et al. (2015). "Toxic Arbitrage." HEC Paris Research Paper No. FIN-2014-1040.
- Glosten, L. R. and P. R. Milgrom (1985). "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." Journal of Financial Economics **14**(1): 71-100.
- Goldstein, M., et al. (2016). "High-frequency Trading Strategies." SSRN Electronic Journal.
- Hagströmer, B. and L. Nordén (2014). "Closing Call Auctions at the Index Futures Market." Journal of Futures Markets **34**(4): 299-319.
- Hasbrouck, J. and G. Saar (2013). "Low-latency Trading." Journal of Financial Markets **16**(4): 664-678.
- Hayes, R., et al. (2012). "Agent Based Model of the E-Mini S&P 500 Future: Application for Policy Making." Computational Intelligence for Financial Engineering & Economics Conference.
- Hendershott, T., et al. (2011). "Does Algorithmic Trading Improve Liquidity?" The Journal of Finance **66**(1): 1-33.
- Hendershott, T. and R. Riordan (2011). Algorithmic Trading and Information. Haas School of Business, University of California at Berkeley, Working Paper.
- Hirschey, N. (2016). "Do High-Frequency Traders Anticipate Buying and Selling Pressure?" SSRN Electronic Journal.
- Huang, R. D. and H. R. Stoll (1997). "The Components of the Bid-Ask Spread: A General Approach." The Review of Financial Studies **10**(4): 995-1034.
- J.P. Morgan (2017). "Big Data and AI Strategies: Machine Learning and Alternative Data Approach to Investing." White Paper: 280.
- Keogh, E. and S. Kasetty (2003). "On the Need for Time Series Data Mining Benchmarks: A Survey and Empirical Demonstration." Data Mining and Knowledge Discovery **7**(4): 349-371.
- Kirilenko, A. A., et al. (2016). "The Flash Crash: the Impact of High Frequency Trading on an Electronic Market." Journal of Finance, Forthcoming.

- Korajczyk, R. A. and D. Murphy (2016). "High Frequency Market Making to Large Institutional Trades." SSRN Electronic Journal.
- Koutrouvelis, I. A. (1980). "Regression-Type Estimation of the Parameters of Stable Laws." Journal of the American Statistical Association **75**(372): 918-928.
- Kyle, A. S. (1985). "Continuous Auctions and Insider Trading." Econometrica **53**(6): 1315-1335.
- Latza, T., et al. (2014). "Fast Aggressive Trading." SSRN Electronic Journal.
- Li, Z. (2015). "Particle Swarm Optimization Using Levy Distribution (Alpha Stable) Randomization." Matlab codes.
- Malinova, K. and A. Park (2016). "'Modern' Market Makers." University of Toronto. Working Paper.
- Malinova, K., et al. (2016). "Taxing High Frequency Market Making: Who Pays the Bill?" SSRN Electronic Journal.
- Menkveld, A. J. (2013). "High Frequency Trading and the New Market Makers." Journal of Financial Markets **16**(4): 712-740.
- Menkveld, A. J. (2016a). "The Economics of High-Frequency Trading: Taking Stock." Annual Review of Financial Economics **8**(1): 1-24.
- Menkveld, A. J. (2016b). "High-Frequency Trading as Viewed Through an Electronic Microscope." SSRN Electronic Journal.
- Menkveld, A. J. and B. Z. Yueshen (2016). "The Flash Crash: A Cautionary Tale about Highly Fragmented Markets." SSRN Electronic Journal.
- Menkveld, A. J. and M. A. Zoican (2017). "Need for Speed? Exchange Latency and Liquidity." The Review of Financial Studies **30**(4): 1188-1228.
- Moshirian, F., et al. (2012). "Overnight Public Information, Order Placement, and Price Discovery during the Pre-opening Period." Journal of Banking & Finance **36**(10): 2837-2851.
- Nolan, J. P. (2015). "Stable Distributions Models for Heavy Tailed Data." American University, Math/Stat department. Working Paper.
- O'Hara, M. (2015). "High Frequency Market Microstructure." Journal of Financial Economics(116): 257-270.
- Paddrik, M. E., et al. (2016). "Visual Analysis to Support Regulators in Electronic Order Book Markets." Environment Systems and Decisions **36**(2): 167-182.
- Pagano, M. S., et al. (2013). "A Call Auction's Impact on Price Formation and Order Routing: Evidence from the NASDAQ Stock Market." Journal of Financial Markets **16**(2): 331-361.
- Pagano, M. S. and R. A. Schwartz (2003). "A Closing Call's Impact on Market Quality at Euronext Paris." Journal of Financial Economics **68**(3): 439-484.

Petitjean, F., et al. (2014). "Dynamic Time Warping Averaging of Time Series Allows Faster and more Accurate Classification." IEEE International Conference on Data Mining, at Shenzhen, China.

Ranaldo, A. (2004). "Order Aggressiveness in Limit Order Book Markets." Journal of Financial Markets **7**(1): 53-74.

Sewell, M. (2011). Characterization of Financial Time Series. UCL Department of Computer Science. Working Paper.

Shorter, G. and R. S. Miller (2014). High-frequency Trading: Background, Concerns, and Regulatory Developments. Congressional Research Service, Washington DC.

Skutkova, H., et al. (2013). "Classification of Genomic Signals using Dynamic Time Warping." BMC Bioinformatics **14** (10).

Stiglitz, J. (2014). "Tapping the Brakes: Are Less Active Markets Safer and Better for the Economy?" Speech at the 2014 Financial Markets Conference at the Federal Reserve Bank of Atlanta.

Subrahmanyam, A. and H. Zheng (2016). "Limit Order Placement by High-Frequency Traders." SSRN Electronic Journal.

Upson, J. and R. A. Van Ness (2017). "Multiple Markets, Algorithmic Trading, and Market Liquidity." Journal of Financial Markets **32**: 49-68.

Vidyasagar, M. (2014). "SYSM 6303: Quantitative Introduction to Risk and Uncertainty in Business Lecture 4: Fitting Data to Distributions." The University of Texas at Dallas, Power Point Presentation.

Yang, L. and H. Zhu (2016). "Back-Running: Seeking and Hiding Fundamental Information in Order Flows." Rothman School of Management Working Paper No. 2583915.

Ye, M., et al. (2013). "The Externalities of High Frequency Trading." SSRN Electronic Journal.

Yergeau, G. (2016). "Profitability and Market Quality of High Frequency Market-Makers: An Empirical Investigation." SSRN Electronic Journal.

Yunyue, Z. and D. Shasha (2003). "Efficient Elastic Burst Detection in Data Streams." Proceedings of the ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: 336-345.

Table 2. 1 Classification of traders

	Position	Inventory	Types
Intraday intermediaries	Small relative to volume	Mean-reverting	HFTers' inventories are positively correlated to stock prices HFMMs' inventories are negatively correlated to stock prices
Fundamental traders	More than 15% of total trading volume	Long Short	Buyer Seller
Small traders	Small	Small if any	Less than 10 contracts per day

This table shows the classification of traders of Kirilenko, Kyle et al. (2016) inferred from the CFTC audit data of the Chicago Mercantile Exchange E-mini S & P 500 futures contract for the period May 3 to 6, 2010. Abbreviations: HFTers: high-frequency traders; HFMMs: high-frequency market makers.

Table 2. 2 Xetra market model

	Opening auction		Intraday auction		Closing auction
DAX	8:50-9:00	Continuous	13:00-13:02	Continuous	17:30-17:35
MDAX	8:50-9:02	trading	13:05-13:07	trading	17:30-17:35

Regular auction periods are followed by a random end of 30 seconds maximum.

Table 2. 3 Auctions: public information

Stock					Indicative	Surplus	Surplus
ID	Date	Time stamp	Status	Matched Q	Price	Ask	Bid
2841	20130201	31800142084	5	20,286	€ 37.925	-	5100
2841	20130201	31800827109	5	25,286	€ 37.925	-	100
2841	20130201	31802184578	5	25,973	€ 37.910	-	413
2841	20130201	31805024087	5	32,291	€ 38.000	-	427
2841	20130201	31805026280	5	32,321	€ 38.100	-	2274
2841	20130201	31810379770	5	33,266	€ 38.200	1,365	-

Stock ID: Stock identifier; Status: 5: opening auction.

Table 2. 4 Interpretation of auction's public information

$\Delta I.Q.$	$\Delta I.P.$	$\Delta \text{Ask Surp.}$	$\Delta \text{Bid Surp.}$	Code	Interpretation
+	+			1	Marketable order buy
+	-			1	Marketable order sell
-	-			2	Cancellation buy
-	+			2	Cancellation sell
0	-	$\neq 0$		1	Sell limit order < indicative price and > best bid
0	+		$\neq 0$	1	Buy limit order > indicative price and < best ask
0	+	$\neq 0$		2	Sell limit order cancellation < indicative price
0	-		$\neq 0$	2	Buy limit order cancellation > indicative price
0	0	$(\Delta \text{ask} - \Delta \text{bid}) > 0$		1	Buy (Sell) limit order if $\Delta \text{bid} > \Delta \text{ask}$ ($\Delta \text{ask} > \Delta \text{bid}$)
0	0	$(\Delta \text{ask} - \Delta \text{bid}) < 0$		2	Cancellation buy (sell) limit order if $\Delta \text{bid} < \Delta \text{ask}$ ($\Delta \text{ask} < \Delta \text{bid}$)

Code 1: creation; Code 2: cancellation; $\Delta I.Q.$: variation in the indicative quantity; $\Delta I.P.$: variation in the indicative price; $\Delta \text{Ask Surp.}$: Variation in the ask surplus; $\Delta \text{Bid Surp.}$: variation in the bid surplus.

Table 2. 5 Behavior-based types of algorithm

	>= 70% at indicative price	< 70% at indicative price
Cancellation >= 70%	Type 6 : Conservative, PDPn, not directionnal	Type 7 : Aggressive, PDP+, not directionnal
Cancellation < 70%	Type 8 : Conservative, PDPn, directionnal	Type 9 : Aggressive, PDP+, directionnal

PDPn: neutral effect on the price discovery process; PDP+: positive effect on the price discovery process.

Table 2. 6 Statistics - all auctions

		#	#	% of	Mean E.O.A.	Median E.O.A.
		Stocks	Events	total events	Matched Q	Matched Q
DAX	Open	30	582,330	17.64%	50,731	18,636
	Midday	30	370,963	11.24%	63,488	8,609
	Close	30	2,347,037	71.12%	264,939	32,484
	# Events	30	3,300,330	100.00%		
MDAX	Open	30	185,020	23.29%	3,235	1,102
	Midday	30	58,384	7.35%	2,273	523
	Close	30	551,017	69.36%	16,935	1,315
	# Events		794,421	100.00%		

Stocks: total number of stocks in the sample; # events: total number of events; Mean E.O.A. Matched Q: Mean end of auction matched quantity; Median E.O.A. Matched Q: Median end of auction matched quantity.

Table 2. 7 Indicative quantity - all auctions

	Variation in Matched Q	# events I.Q.	% events I.Q.	Mean I.Q.	Median I.Q.
DAX	= 0	1,741,847	52.78%		
	> 0	496,199	15.03%	3,815	430
	< 0	1,062,314	32.19%	- 3,061	- 302
	Total	3,300,360	100.00%		
MDAX	= 0	405,631	51.06%		
	> 0	282,214	35.52%	764	152
	< 0	106,576	13.42%	-524	-136
	Total	794,421	100.00%		

events I.Q.: number of events affecting the indicative quantity; % events I.A.: percentage of total number of events for the category; Mean I.Q.: average indicative quantity; Median I.Q.: median indicative quantity.

Table 2. 8 NPLs algorithm types

Type of algorithm	Matched quantity	Surplus bid	Surplus ask
1	0	0	< >
2	0	< >	0
3	< >	< >	0
4	< >	0	< >
5	< >	< >	< >

< > : a positive or negative variation in the variable; 0: no impact.

Table 2. 9 Negative price loops and SWT levels

SWT levels	2	3	4	5	NPL	All	% all
Open	170	268	788	-	1,226	582,330	0.21%
Midday	-	-	9	11	20	370,963	0.01%
Close	248,974	56,826	96,700	-	402,500	2,347,037	17.15%
NPL	249,144	57,094	97,497	11	403,746	3,300,330	12.23%
% by level	61.71%	14.14%	24.15%	0.00%			

SWT levels: shifted wavelet tree level at which algorithmic sequences have been identified. A measure of latency.

Table 2. 10 NPLs algorithmic sequences

DAX		Total	Open	Midday	Close
Sequences		2,595	79	2	2,514
Stocks		30/30	17/30	2/30	30/30
Repetition	Mean	143	9	8	147
	Median	61	8	8	65
Duration (sec.)	Mean	21.150	0.479	0.169	21.818
	Median	2.230	0.097	0.169	2.422
Duration (sec.)	Mean	0.460	0.040	0.022	0.473
Between events	Median	0.020	0.012	0.022	0.020
MDAX					
Sequences		134	1	-	133
Stocks		8/30	1/30	-	7/30
Repetition	Mean	355	52	-	357
	Median	144	52	-	144
Duration (sec.)	Mean	22.656	2.108	-	22.719
	Median	3.535	2.108	-	3.616
Duration (sec.)	Mean	0.186	0.041	-	0.187
Between events	Median	0.018	0.041	-	0.018

Sequences: algorithmic sequences from NPL; Repetition: Number of repetitions during the sequence; Duration (sec.): duration in second of a sequence; Duration (sec.) mean, median: average (median) duration between events in a sequence.

Table 2. 11 Stable family parameters February – July 2013

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	alpha			beta			gamma		delta	
id	w/o	with	diff (%)	w/o	with	diff (%)	w/o	with	w/o	with
22	0.974	0.886	9.02%	-0.143	-0.1349	5.40%	2,908	2,419	9,684	1,545
24	1.053	0.985	6.48%	-0.113	-0.1265	-12.35%	1,808	1,585	- 2,271	8,713
32	0.944	0.582	38.39%	-0.084	-0.2422	-189.94%	1,304	644	636	170
49	1.284	1.209	5.79%	-0.078	-0.0365	53.09%	3,892	3,630	- 1,520	- 1,154
58	0.827	0.690	16.55%	-0.018	-0.0007	95.93%	1,515	1,086	97	- 20
60	0.826	0.721	12.72%	0.055	0.1207	-120.08%	1,314	1,023	- 110	- 222
80	1.118	0.965	13.63%	-0.001	0.0466	3328.68%	1,857	1,482	39	- 1,251
85	0.915	0.844	7.69%	-0.127	-0.0161	87.31%	1,089	970	1,163	94
106	1.072	1.042	2.78%	-0.030	0.4028	1440.98%	5,803	5,458	- 1,356	30,727
130	1.100	1.046	4.87%	-0.053	-0.0138	74.00%	1,846	1,678	- 491	- 224
138	0.600	0.586	2.31%	0.095	0.0204	78.41%	52,972	49,900	- 2,535	333
143	0.875	0.815	6.85%	-0.056	-0.0435	22.79%	7,790	6,698	1,848	644
146	1.139	1.083	4.91%	-0.017	-0.0430	-154.55%	3,797	3,530	- 338	- 1,290
151	0.969	0.931	3.96%	-0.054	0.0090	116.69%	1,734	1,569	2,206	94
266	0.493	0.410	16.73%	-0.039	-0.1230	-219.35%	1,335	964	8	14
829	0.729	0.486	33.38%	-0.113	-0.0819	27.33%	10,454	5,749	2,731	78
1634	0.914	0.817	10.63%	0.003	-0.0171	711.05%	1,233	985	- 12	67
2451	0.978	0.941	3.80%	-0.175	-0.1608	8.36%	5,222	4,511	25,654	7,189
2481	0.764	0.481	36.98%	-0.268	-0.2975	-10.93%	6,210	3,465	3,248	407
2807	1.195	1.157	3.23%	-0.061	-0.0233	62.07%	3,900	3,659	- 593	- 199
2841	1.045	0.954	8.73%	-0.546	-0.2626	51.94%	7,534	6,787	- 54,551	25,411
3446	0.873	0.737	15.51%	-0.058	-0.0517	11.05%	1,847	1,569	718	267
3679	0.956	0.880	7.92%	-0.076	-0.0227	70.14%	23,884	19,706	25,938	1,572
3744	1.266	1.150	9.13%	-0.076	-0.0297	60.95%	25,672	22,063	- 3,030	- 2,053
4423	0.710	0.523	26.26%	-0.103	-0.0421	59.03%	1,350	883	352	- 4
5830	0.955	0.909	4.83%	0.013	0.0587	-347.29%	2,650	2,308	- 260	- 803
8669	0.455	0.414	8.98%	-0.008	0.0203	349.50%	864	551	51	- 31
9633	1.072	0.994	7.25%	-0.047	0.0000	100.00%	3,751	3,167	- 1,799	- 283
11814	0.977	0.828	15.24%	-0.010	0.0709	820.50%	9,623	7,215	4,148	- 1,112
16753	0.984	0.915	6.97%	0.112	0.0965	13.58%	5,310	4,704	- 22,055	- 2,475

$\lambda \in (0,2]$: stability parameter; $\beta \in [-1,1]$: skewness parameter; $\gamma \in (0,\infty)$: scale parameter; $\delta \in (-\infty, +\infty)$: location parameter.

Table 2. 12 Kolmogorov-Smirnov, two-sample tests

Stock			Stock		
id	p-value	k-s stat	id	p-value	k-s stat
22	0.00000	0.06491	829	0.00000	0.05155
24	0.00000	0.09367	1634	0.00000	0.02561
32	0.00000	0.22691	2451	0.00000	0.03294
49	0.00001	0.01345	2481	0.00000	0.08661
58	0.00000	0.03625	2807	0.00000	0.01545
60	0.00000	0.07546	2841	0.00000	0.09443
80	0.00000	0.03203	3446	0.00000	0.03616
85	0.00000	0.06835	3679	0.00784	0.00861
106	0.00000	0.12032	3744	0.00000	0.01923
130	0.00000	0.02870	4423	0.00000	0.07555
138	0.00000	0.04474	5830	0.00000	0.01706
143	0.00002	0.01547	8669	0.00000	0.07345
146	0.00000	0.03958	9633	0.00000	0.06727
151	0.00000	0.03059	11814	0.00000	0.02252
266	0.00000	0.04727	16753	0.00000	0.05977

p-value: Kolmogorov-Smirnov p-value; k-s stat: Kolmogorov-Smirnov statistic.

Table 2. 13 Behavior-based algorithmic events

		#	#	%
		events	identified	identified
DAX	Open	582,330	72,514	12.45%
	Midday	370,963	53,023	14.29%
	Close	2,347,037	332,327	14.16%
		3,300,330	457,864	13.87%
MDAX	Open	185,020	6,106	3.30%
	Midday	58,384	4,677	8.01%
	Close	551,017	127,573	23.15%
		794,421	138,356	17.42%
Total		4,094,751	596,220	14.56%

events: total number of auctions events; # identified: number of events identified as algorithmic sequences;
 % identified: total events' percentage due to identified algorithms.

Table 2. 14 Behavior-based types of algorithm

	Algorithm	#	%
	Type	identified	identified
DAX	6	293,042	64.00%
	7	31,970	6.98%
	8	119,740	26.15%
	9	13,112	2.86%
	Total	457,864	100.00%
MDAX	6	9,319	6.74%
	7	1,966	1.42%
	8	126,463	91.40%
	9	608	0.44%
	Total	138,356	100.00%

Algorithms characteristics: Type 6: conservative, neutral effect on the price discovery process, positions not directional; type 7: aggressive, positive impact on the price discovery process, positions not directional; type 8: conservative, neutral effect on the price discovery process, positions directional; type 9: aggressive, neutral effect on the price discovery process, positions directional; # identified: number of events identified as algorithmic sequences; % identified: total events' percentage due to identified algorithms.

**Table 2. 15 DAX Behavior-based sequences : statistics by auction
February – July 2013**

	Type	# orders	% identified	# stocks	% at i.p.
Open	6	41,032	56.58%	30	93.23%
	7	4,926	6.79%	30	93.26%
	8	24,212	33.39%	30	27.42%
	9	2,344	3.23%	30	31.78%
Midday	6	21,062	39.72%	30	93.56%
	7	2,834	5.34%	30	90.08%
	8	27,887	52.59%	30	14.86%
	9	1,240	2.34%	29	49.68%
Close	6	230,948	69.49%	30	98.16%
	7	24,210	7.28%	30	95.14%
	8	67,641	20.35%	30	10.13%
	9	9,528	2.87%	30	39.35%
	Open	72,514	15.84%		
	Midday	53,023	11.58%		
	Close	332,327	72.58%		
	ALL	457,864	100.00%		

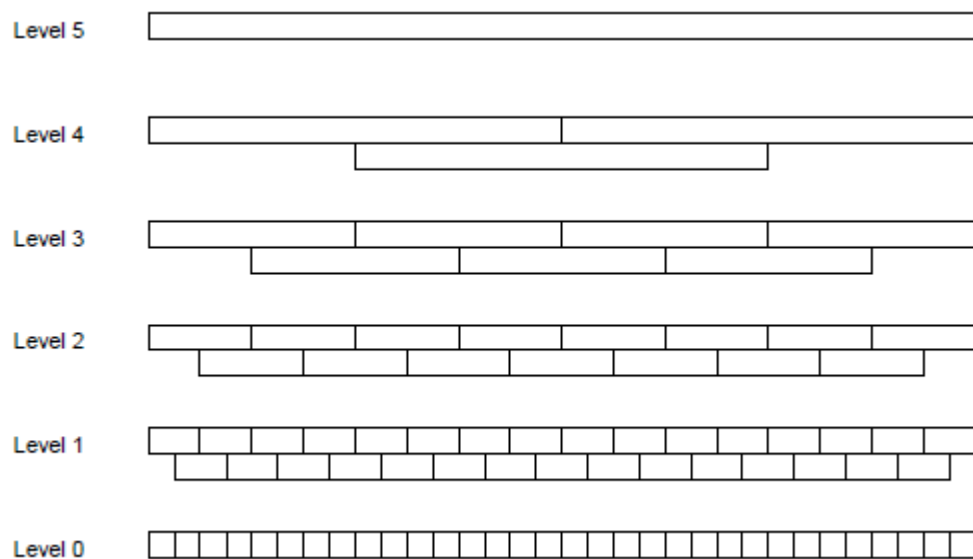
Type: 6: conservative, neutral effect on the price discovery process, not directional; type 7: aggressive, positive effect on the price discovery process, not directional; type 8: conservative, neutral on the price discovery process, directional; type 9: aggressive, positive on the price discovery process, directional; # orders: total number of orders (creation and cancellation); % identified: percentage of identified sequences of type x; # stock: number of stocks with at least one algorithmic sequence; % at i.p.: percentage of orders at the indicative price.

**Table 2. 16 MDAX Behavior-based sequences : statistics by auction
February – July 2013**

MDAX			%	#	%
	Algo.	# orders	identified	stocks	at i.p.
Open	6	2,807	45.97%	26	99.00%
	7	511	8.37%	29	96.67%
	8	2,645	43.32%	20	10.78%
	9	143	2.34%	24	39.16%
Midday	6	4,455	95.25%	5	99.98%
	7	88	1.88%	13	94.32%
	8	119	2.54%	6	22.69%
	9	15	0.32%	2	53.33%
Close	6	2,057	1.61%	19	99.37%
	7	1,367	1.07%	28	96.34%
	8	123,699	96.96%	22	0.14%
	9	450	0.35%	20	30.22%
	Open	6,106	4.41%		
	Midday	4,677	3.38%		
	Close	127,573	92.21%		
	ALL	138,356	100.00%		

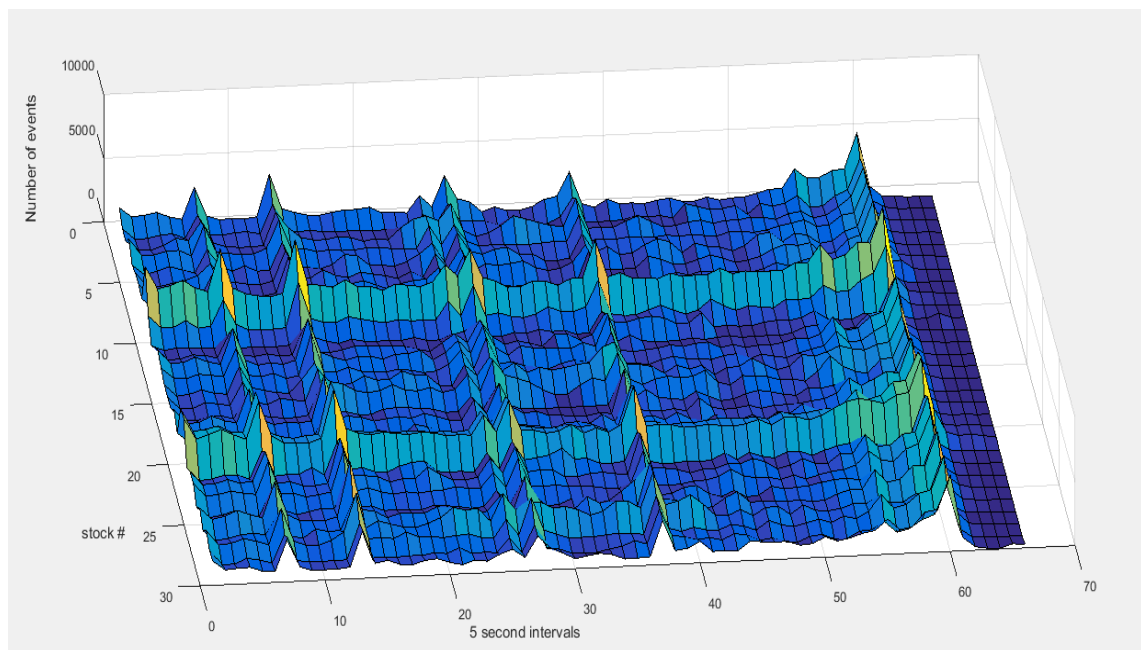
Type: 6: conservative, neutral effect on the price discovery process, not directional; type 7: aggressive, positive effect on the price discovery process, not directional; type 8: conservative, neutral on the price discovery process, directional; type 9: aggressive, positive on the price discovery process, directional; # orders: total number of orders (creation and cancellation); % identified: percentage of identified sequences of type x; # stock: number of stocks with at least one algorithmic sequence; % at i.p.: percentage of orders at the indicative price.

Figure 2. 1 Shifted wavelet tree - structure



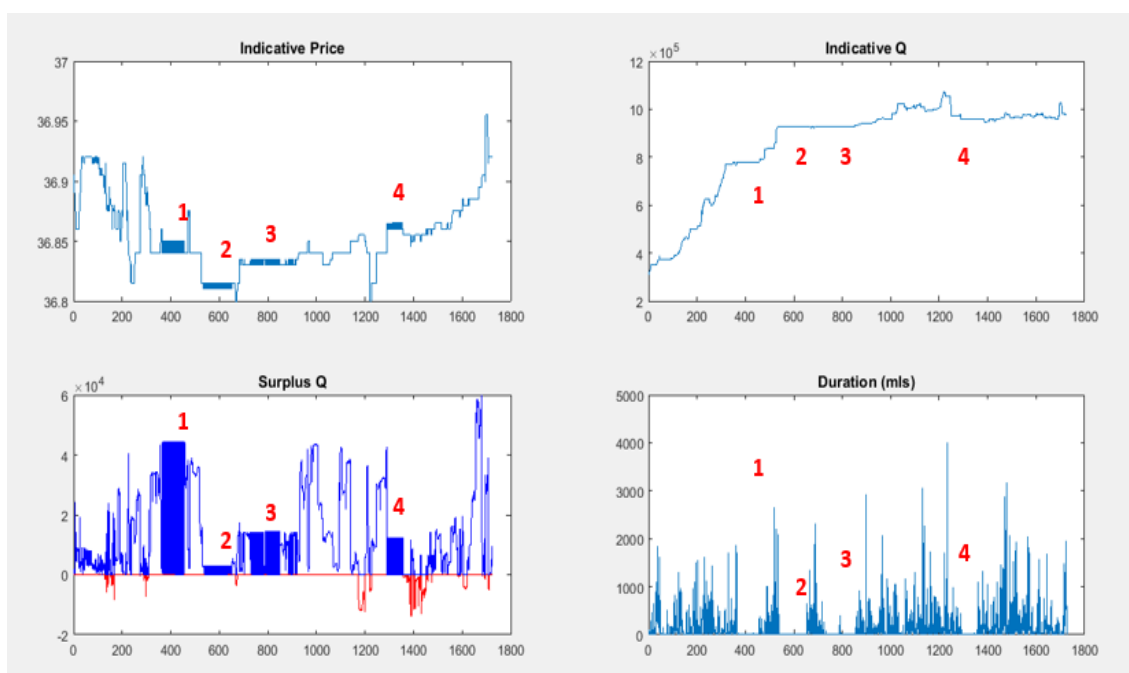
A graphical representation of SWT data structure. Level 0 contains aggregates from the smallest discrete time interval. Level 1 highest row is the pairwise sum of each level 0 two consecutive data, starting with the first time interval of level 0. The level 1 lowest row is the pairwise sum of each level 0 two consecutive data, starting with the second time interval of level 0. This creates the observed overlapping. Process is repeated for higher levels.

Figure 2. 2 DAX components - closing auctions



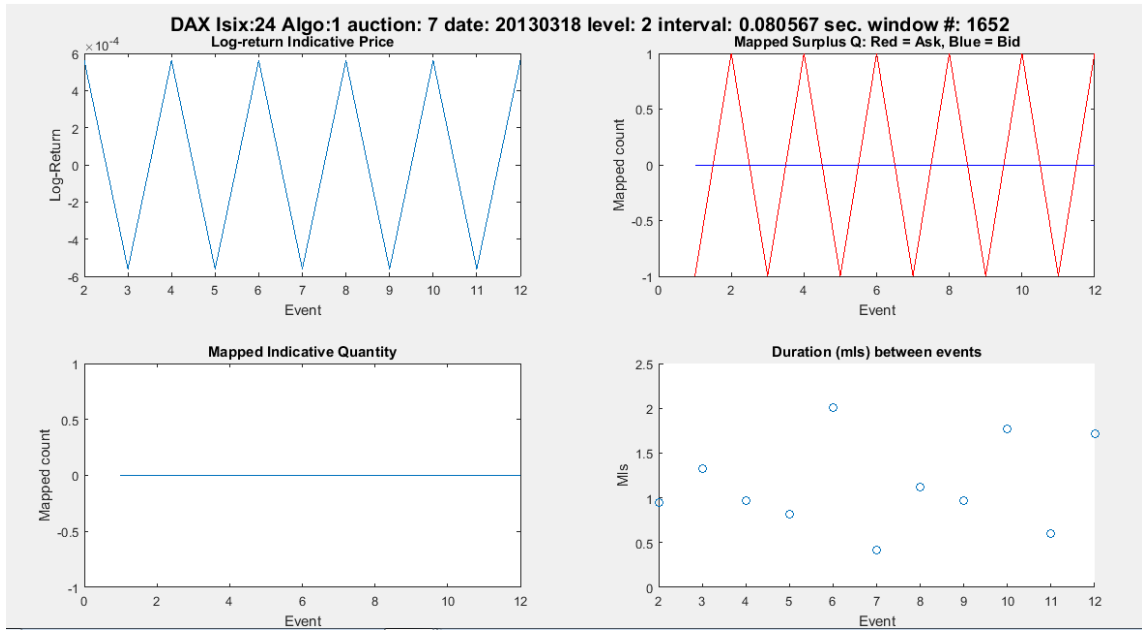
Cyclicality in the number of events affecting all stocks: institutional investors imprints (Easley, Lopez de Prado et al. (2012)).

Figure 2. 3 Deutsche Bank 20130212 closing auction



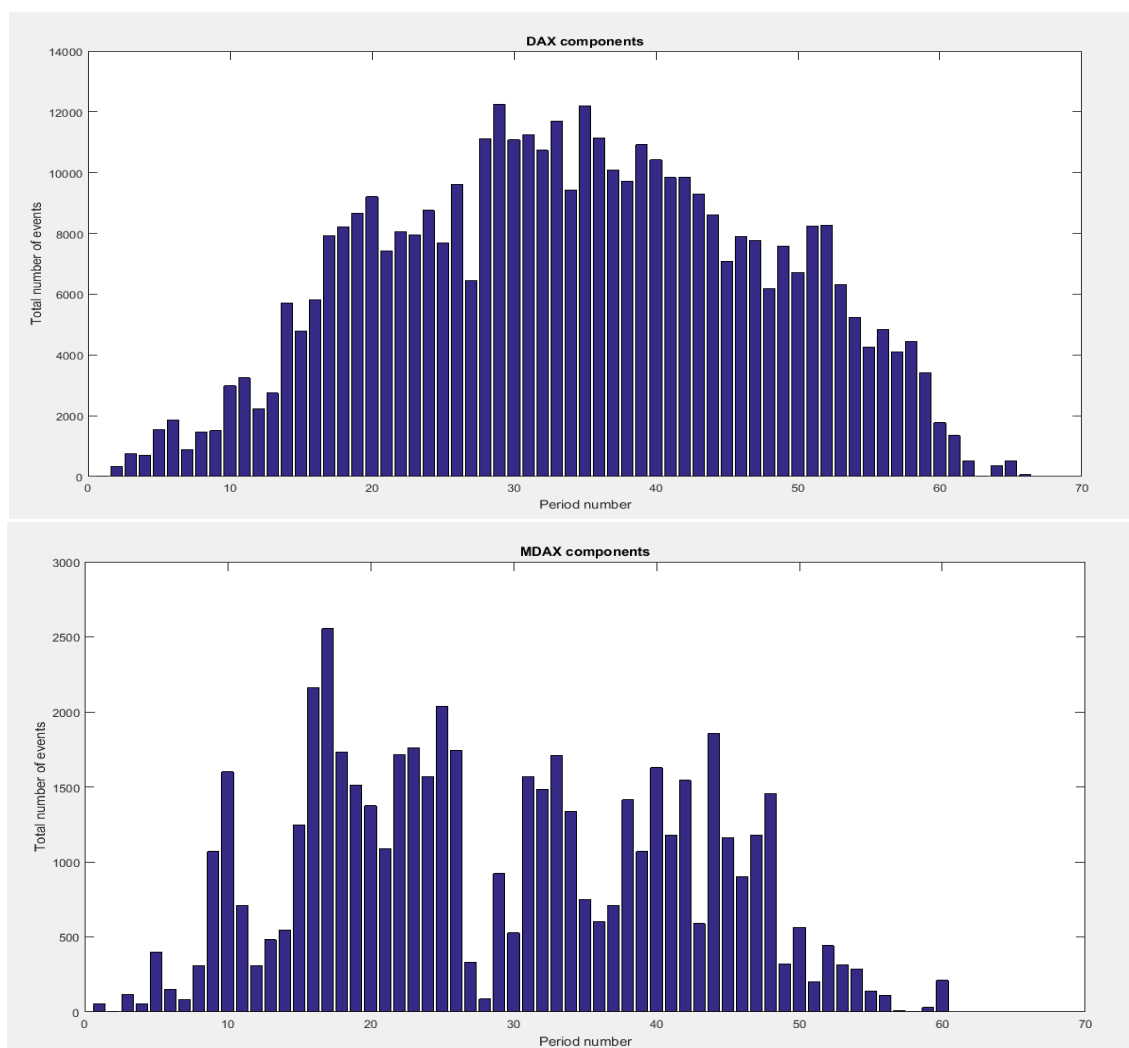
Indicative Q: matching quantity at the indicative price; Indicative price: price maximizing the matching quantity; Surplus Q: imbalance; Duration: time lapse between two events; A graphical representation of four NPLs labeled from 1 to 4. We comment the graphics clockwise. Indicative price's volatility changes with sequences. They influence marginally the indicative quantities. NPLs' aggressiveness and/or LOB depth have significantly different impacts on imbalances (surplus quantities). All NPLs are executed via ultra high-frequency algorithms.

Figure 2. 4 Identification parameters for negative price loops



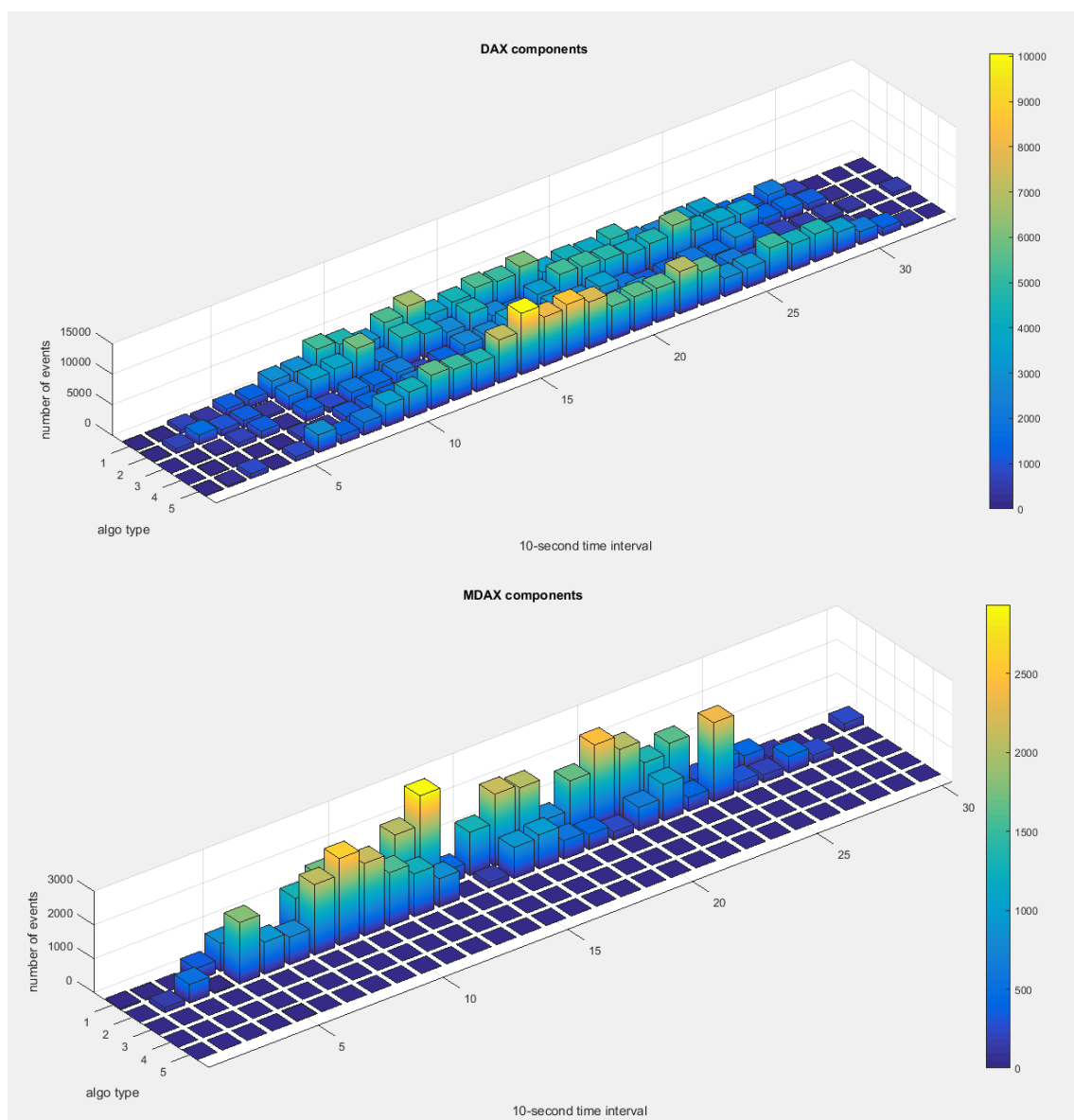
Mapped surplus $q \in (-1,1)$; mapped indicative $q \in (-1,1)$; 0: no variation in mapped variable; mls: milisecond; Following a burst indicator's alarm from a SWT tree, characteristics from a potential NPL are graphed. Events come from a time interval of 0.08 second (SWT level 2). We interpret clockwise. The indicative price's logarithmic returns and the surplus quantities vary symmetrically confirming the NPLs' creations-cancellations sequence characteristics. Limit orders involved are inside the bid-ask spread on the offer side as matched (indicative) quantities do not fluctuate. Duration requires high-frequency technologies. DTW compares the sequence to pre-identified ones and diagnoses NPL existence.

**Figure 2. 5 NPL events by 5-second interval - closing auctions :
aggregated statistics February – July 2013**



Period number: number of 5-second periods elapsed since the beginning of the auction; total number of events: index's total number of events for a given 5-second interval.

Figure 2. 6 NPL events by 10-second interval – closing auctions : type of algorithms February – July 2013



x axis: number of 10-second time intervals elapsed since the beginning of the auction; y axis: algorithmic sequence type; z axis: total number of algorithmic imprints occurrences.

Figure 2. 7 Transaction cost of NPL sequences

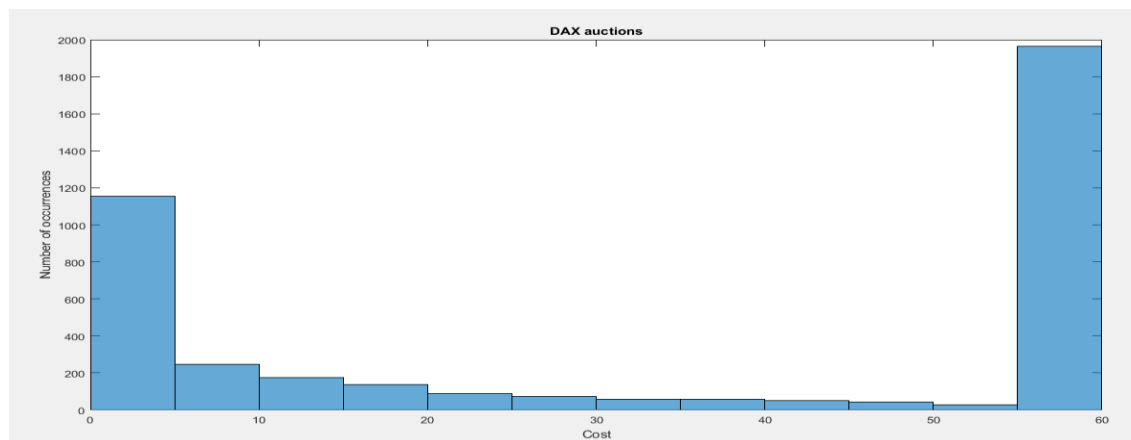
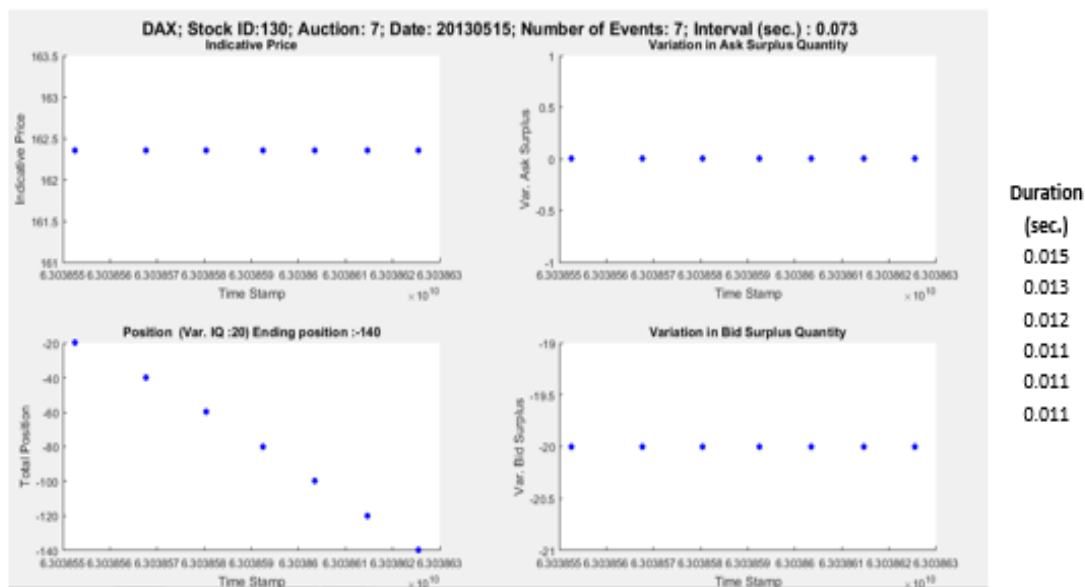


Figure 2. 8 TWAP example - Volkswagen



Stock id: unique stock identifier; auction=7: closing auction; TWAP: time weighted average price algorithm.

Chapter 3

Performance and behavior of endogenous liquidity providers

Abstract

We evaluate the impact of broker preference, a Canadian market microstructure specificity, on the performance of endogenous liquidity providers (ELPs). Using the international standard of price-time prioritization as a benchmark and the characteristics of two brokers with a significantly different clientele, we analyze the performance of ELPs with a parsimonious model that avoids data mining, considers transaction fees and rebates, and accounts for market imperfections. Special care is given to respect the price and quotes discovery processes. Profitability is high, even excluding discounts to liquidity providers. P&Ls volatility is low and no significant drawdown occurs.

The ELPs positions are positively correlated with the intensity of trades and quotes and the bid-ask spread. Increases in volatility create an incentive to participation, a sign of ELP's presence in volatile environments. Liquidity imbalances and the market momentum do not significantly influence their positions. We analyze the behavior of ELPs in extreme situations. Although a detailed analysis of critical events shows that the ELPs did not withdraw from the market during the period covered by our sample, we cannot exclude illiquidity contagion in the future.

Introduction

Broker preference is a method of prioritizing the allocation of marketable orders unique to Canada. At a given price, the Stock Exchange matches the incoming marketable orders identified to a broker with that broker's limit orders, notwithstanding its time priority in the limit order book (LOB). We assess the impact of this particularity by comparing the performances of endogenous liquidity providers using two scenarios. First, we apply the price-time prioritization international standard. Second, we enforce broker preference and we trade successively as two brokers having a significantly different clientele basis. In order to assess the economic impact of these prioritization scenarios, an endogenous liquidity provider (ELP) having as main activity to supply liquidity, uses an algorithm

conceptualized by Yergeau (2016). The ELP is not obligated to supply liquidity or otherwise facilitate trading. Quoting decisions come from Ait-Sahalia and Saglam (2014) dynamic inventory management model. We embed the model in a strategy that includes circuit breakers, considers market imperfections and closes positions at the end of the day. No trade nor quote prices are modified by the ELP. This reflects a preoccupation to respect the trade and quote discovery processes as realised. We estimate ex-ante the Ait-Sahalia and Saglam (2014) model parameters with a methodology applied uniformly to all stocks. As no data mining is used, we limit the possibility of false discovery (Lopez De Prado (2015)). Simulations are possible through the use of a proprietary database containing identifiers for all creations, cancellations, modifications, and executions of limit and marketable orders received by the Stock Exchange on the S&P/TSX 60 components between March 1, 2015 and August 31, 2015. This information allows to follow all orders life cycle.

ELPs play an important role in stock markets (Malinova, Park et al. (2016)). Their importance is due to the development in information technology and direct market access (DMA) that allow anyone to act as a liquidity provider (Baldauf and Mollner (2016)). ELPs' economic viability is *sine qua non* for effective financial intermediation, since without this profitability they withdraw. Despite their importance, few studies assess the viability of their operations. A new market maker on Chi-X Europe generated an average profit of € 9,500 per day between January 1, 2007, and June 17, 2008 (Menkveld (2013)). Yergeau (2016) simulates the activities of an ELP who trades 60 stocks from the Frankfurt Stock Exchange during March to July 2013. The ELP operations lead to a net profit of 3 million euros. Other studies target only a portion of liquidity provider operations and have mixed results. Korajczyk and Murphy (2016) focus on the profitability of ELPs facing large orders issued by institutional investors on the Toronto Stock Exchange (TSX) from January 1, 2012, to June 30, 2013. Economic viability is exclusively due to the collection of liquidity discounts from the Stock Exchange's maker-taker model. On the American side, Sofianos and Xiang (2013) analyze Goldman Sachs's agency trades. They pinpoint that high-frequency traders struggle to make their operations profitable when they act as counterparts. Brogaard, Hendershott et al. (2014) delve into the aggregated performance

of high-frequency traders. They find that their activities in large-cap stocks are lucrative.⁸ Our analysis contributes to the existing literature as we evaluate the ELPs viability. In all scenarios, profits are economically significant without the rebates and justify the algorithm implementation. No significant drawdown occurred. Under broker preference, the two brokers obtain trading profits appreciably different. The broker with the highest customer base earns lower trading profits of \$ 553k over the six-month period. The maker-taker model compensates for this by equating the performances of both brokers. If we add trading fees charged to their respective clientele, the most active broker obtains the highest net revenue. The algorithm respects the stylized facts about the high-frequency market makers behavior. We identify the conditions that influence the ELPs positions. The ELPs react to variation in trade and quotes intensities and bid-ask spreads. Their positions are positively correlated with stocks' idiosyncratic risk and are not influenced by variations in liquidity imbalances and market momentum.

Even in the presence of profitability, the absence of formal obligations to provide liquidity is of concern to both the regulators (U.S. Securities and Exchange Commission (2010); Basel Committee on Banking Supervision (2013)), and the institutional investors (Hope (2013); Swedroe (2016)). The main discomforts are related to the coordinated fluctuations in ELPs activities, a potential cause of systemic illiquidity. The same worries apply on the Canadian side. IIROC gave access to audit databases to investigate ELPs behaviors. Korajczyk and Murphy (2016) note that HFT liquidity provision is significantly reduced when facing large trades. Anand and Venkataraman (2016) analyze the covariations of the liquidity supply of ELPs. They observe a synchronous shrinkage that periodically decreases liquidity. We analyze the illiquidity contagion risk in extreme situations by examining the cross-correlation coefficients of ELP's relative participation to the bid and offer at LOB level 1. The cross-correlation matrix distribution is centered around a value slightly below +0.20, which implies positively correlated behavior of ELP's liquidity from both sides of the market. Detailed analysis of critical situations reveals eight time intervals where the ELP has completely stopped providing liquidity. Of these, seven occur

⁸ Chung, K. H. and A. J. Lee (2016). "High-frequency Trading: Review of the Literature and Regulatory Initiatives around the World." *Asia-Pacific Journal of Financial Studies* **45**(1): 7-33.
 present a literature review of the aggregate profitability from the high-frequency trading industry.

in the last two minutes of trading and are due to a risk management feature closing positions. The last one occurred on August 25, 2015, at 10:27. It cannot be dissociated from the fundamental information content that caused a starting market downturn.

The paper is organized as follows. Section 3.1 describes the characteristics of trading on the TSX. Section 3.2 presents the data. Section 3.3 describes the functioning of the algorithm. Section 3.4 presents the methodology applied to the simulation of ELP activities. Section 3.4 discusses the results. Section 3.6 concludes.

3.1 Trading on the TSX

Per the Investment Industry Regulatory Organization of Canada (IIROC), the Canadian stock market structure has twelve entities active in January 2015. Of the total \$ 220.5 billion traded in the country, 56.6% is traded on the Toronto Stock Exchange (TSX). Their closest competitor is the Nasdaq CXC with 16.4% (IIROC (2016)). Despite fragmentation, financial intermediation remains focused on the TSX. Trading on the TSX takes place over the four sessions described in Table 3.1. The market on close session overlaps the session in continuous time and its activities are handled separately. We focus on the continuous time session of the TSX which represents more than 95% of the activities. Every stock has a designated market maker (DMM). The DMM ensures a fair and orderly market by maintaining price continuity and reasonable liquidity. The DMM is responsible for quoting on both sides of the market according to volume and spread requirements. Concurrently, ELPs can participate in the LOB.

[insert Table 3.1 here]

The TSX enforces a maker-taker fee model for more than a decade. It was implemented to encourage and reward liquidity provision (TMX Group (2015)). The maker-taker model has been modified on June 1, 2015 (see Table 3.2). It resulted in a 38.7% (34.4%) decrease in the maker rebate (taker fee). IIROC imposes monitoring fees based on the number of order creations and cancellations at \$ 0.00022 per unit (Malinova, Park et al. (2016); IIROC (2014)).

[insert Table 3.2 here]

3.2 Data

We use proprietary data from a Canadian bank which includes the S&P/TSX 60 components. These stocks have the highest market capitalization. Our sample encompasses from March 2015 to August 2015 inclusively. All limit and marketable orders are available for the four trading sessions of Table 3.1. All orders have a unique identifier (ID). It is therefore possible to monitor their evolution over time (creation, cancellation, modification and execution). The identification of brokers issuing limit orders and brokers on both sides of marketable orders is available but not the traders ID. The order and brokers IDs lead to a LOB's reconstruction which allows to evaluate the impact of different prioritization scenarios. The LOB's reconstruction is possible thanks to Bilodeau (2016)⁹. Seven dates were excluded from the sample due to problems associated with the processing of messages issued by the Stock Exchange¹⁰. Events are timestamped in microseconds.

The six major Canadian banks dominate trading on the TSX. They account for more than 50% of all trades. Brokers have the option of not revealing their identity by codifying the orders as anonymous. In this case, price-time prioritization applies (TMX Group (2015)). Brokers with a large customer base tend not to use anonymous orders to benefit from broker preference (Anand and Venkataraman (2016)). Other brokers could use anonymous orders to conceal their intentions.

Table 3.3 summarizes the monthly S&P/TSX 60 activities between March and August 2015. The stocks are liquid: more than 35 million trades (Col. 3) and 540 million quotes on both sides of the LOB first level (Col. 4) took place. A regime shift in the quote to trade ratio (Col. 5) occurred in June. This regime shift involving a sharp decline in the ratio coincided to the modification of the maker-taker model (Table 3.2). The depth of the market (Col. 7) and the number of orders at the LOB's first level (Col. 8) remained constant.

⁹ Yann Bilodeau developed TSXParser to process data from the Toronto Stock Exchange.

¹⁰ The missing dates are March 31, May 6, May 7, May 15, May 28, May 29, and July 16.

[insert Table 3.3 here]

Traders interest decreased from March to May: the market values (Col. 1) and the number of trades (Col. 3) decreased significantly. Adjustments to the statistics for the missing dates mitigate these effects. For example, in May the adjusted number of trades is 4.9 m or 83.5% of the observed average. Activity returned in June and peak in August 2015 with more than 50.5 billion \$ of market values traded (Col. 1). Many factors led to the beginning of a market correction of August 2015: the long-term monetary deficit in Greece, the slowdown in Chinese economic activity, and concerns about the prospect of an imminent end to the US Federal Reserve's quantitative easing policy. These factors led to the collapse of stock markets around the world (Shan Li, Chang et al. (2015); Irwin (2015)). Figure 3.1 illustrates the impact of the onset of this crisis on the S&P/TSX 60 VIX Futures, a market risk indicator based on the intraday price of two S&P/TSX 60 Index short-term options (S&P Indices (2010)). From an average smaller than 15 from March up to 20 August 2015, it rises to an average of 27 from August 21 to August 31.

[insert Figure 3.1 here]

3.3 Endogenous liquidity provider

Our ELPs utilize the optimal quoting policies (OQPs) of Ait-Sahalia and Saglam (2014)'s dynamic inventory management model to decide whether or not to provide liquidity on one or both sides of the LOB's first level. An algorithm embeds the model. It includes circuit breakers whose importance is documented in the literature (Jain, Jain et al. (2016); Chung and Lee (2016); Black Rock (2015); The Government Office for Science London (2012)). The algorithm uses as a risk management feature a uniform closing procedure at the end of the day. The simulations consider market imperfections and the specificities of the Stock Exchange. The latency, the time required to receive, process and respond to new information, is 50 microseconds.

3.3.1 Optimal quoting policy and circuit-breakers

OQP's estimation procedures are presented in Appendix 3.1. Succinctly, we define the ex-ante values of the six model parameters before the simulations. The parameters D , the

discount rate, Γ , the inventory aversion coefficient, and P , a measure of signal quality, apply uniformly to all stocks. In Ait-Sahalia and Saglam (2014) model, market states governing the OQPs are conditioned on parameters λ , the Poisson distribution arrival rate of the incoming low-frequency traders market orders, μ , the Poisson distribution arrival rate of the ELP's signals, and C , the bid-ask spread. We obtain the parameters value from one-minute time interval activities using the first five days of March 2015. Results are presented in Tables A3.2. 1 (λ), A3.2. 2 (μ) and A3.2. 3 (C) of Appendix 3.2. The algorithm uses the thresholds of an OQP (the decisions to quote or not and the quantities implied) as long as the values for every of the parameters λ , μ and C are equal to or less than their respective quantile values of 95%. When a parameter exceeds this quantile value, the algorithm activates a circuit breaker: the ELP cancels its limit orders and liquidates its position conditionally to the provisions set out in Section 0. The ex-ante OQPs are obtained from Equation (3.1).

$$\begin{aligned} \min_j &= \text{quantile}_j(95\%)/4; \text{step}_j = \text{quantile}_j(95\%)/4; \max_j \\ &= \text{quantile}_j(95\%), \end{aligned} \quad (3.1)$$

where:

\min_j : Minimum value of parameter j during interval t .

quantile_j : Quantile of the parameter j during interval t .

step_j : Value of the variation of the parameter j during interval t .

\max_j : Maximum value of parameter j during interval t .

For an ELP to quote on a stock while using Ait-Sahalia and Saglam (2014) model, the combination of parameters λ , μ and C must obtain a positive expected reward. This is the case for all stocks except CSU. Intuitively, 10 trades per minute (Table A1.1, col. 7) seems insufficient to compensate the market risk inherent to intraday positions.

During the simulations, the algorithm monitors market conditions and adjusts the OQPs accordingly. Using ABX as an example, the values of 81 trades (Table A1.1, Col.7), 1,172 signals (Table A1.2, Col.7) and a bid ask-spread of \$ 0.02 (Table A1.3, Col.8) are the OQPs parameters' upper limits. If the number of ABX trades exceeds 81 during a given time interval, the algorithm activates a circuit breaker which cancels all ABX limit orders

and issues a marketable order to liquidate the existing position. Regular quoting activities are induced when parameter values are reset. Chordia, Goyal et al. (2013) note that market makers are also liquidity takers in their regular activities.

3.3.2 Closing of positions and market imperfections

The algorithm closes positions at the end of the day by issuing marketable orders from the last three minutes of trading. This is consistent with the definition of high-frequency traders from the U.S. Securities and Exchange Commission (2010). IIROC (2014) also considers inventory management dynamics and net positions in the traders classification. These closures ensure that ELP's performance is as close as possible to the activities of a liquidity provider as they avoid directional positions after the market close.

Empirical verification is subject to bias when they fail to comply with the price discovery process and the LOB's activities as they occur. We compel the ELP to abide by them: it limits itself to react to marketable and limit orders modifying the level 1 prices generated by the other participants. If the liquidation of a position induced by a circuit breaker or the end-of-day procedure involves walking in the LOB, the ELP trades up to the quantity available at the first level. If required, the ELP waits for the next limit order to trade at level 1 in order to completely liquidate its position. This is tantamount to controlling the instantaneous impact on prices (Cont, Kukanov et al. (2014); Bouchaud, Farmer et al. (2009)). The permanent impact of the ELP's trades and quotes (Hautsch and Huang (2012); Huh (2014); Zhou (2012)) is in this way mitigated.

3.3.3 Scenarios

The simulations quantify the impact on performance of two prioritization scenarios. The first scenario applies price-time priority, the international standard. The second scenario applies the Canadian prioritization: when a marketable order originates from a reference broker, its counterpart is allocated to the reference broker's limit orders according to their time priorities. If the marketable order's quantity exceeds the total quantity of the reference broker's limit order(s), the excess quantity is allocated according to time priority of the other LOB's limit orders.

3.4 Simulation of activities

To limit the probability of false discoveries, we use a single methodology designed *ex ante*. Continuous time session begins with the transfer of orders not executed during the market on open session defined in Table 3.1. The Stock Exchange adds the unrestricted limit orders from previous days and those created at the beginning of the current session. The simulation reproduces the arrival of limit and marketable orders. Each stock can have different OQPs during a given time interval: during the asynchronous arrivals of information of a time interval, the OQPs are based on the cumulated values of the parameters λ and μ , and the observed value of parameter C . The OQP corresponding to the combination of the contemporaneous parameters generates the decisions to provide liquidity or not and the quantities to be quoted.

The arrival of a marketable order entails the reconstruction of the LOB's first level to the opposite side to the arrival of the marketable order. One of the prioritization scenarios from Section 3.3.3 (price-time or broker preference) determines the ELP's participation to the new marketable order. This induces a dynamic position management due to variations in the quantities of the marketable orders and of the OQP's bid and offer limit orders. A circuit breaker triggering or the third minute before the end of the session cause the cancellation of the ELP's limit orders and the liquidation of its positions conditionally to the respect of the price discovery process and of the quotes in accordance to the provisions of Section 3.3.2.

3.5 Results

The cornerstone of the ELP's activities is derived from the OQPs determined by the Ait-Sahalia and Saglam (2014) model. This model formalizes the behavior of a single stock and does not take into account the effects of diversification and arbitrage (spatial and cross-stocks). Simulations respect these two constraints: decisions about one stock are independent from the positions and limit orders of the other stocks from the sample.

The Canadian and US stock markets share many characteristics. Their volatility is strongly correlated (Malinova, Park et al. (2014)). The average trading cost is roughly

similar in both markets (Frijns, Indriawan et al. (2014)). More than 2/3 of the S&P/TSX 60 components trade in US markets. US activities account for approximately 50% of the total trade volume of the Canadian constituents of the index (Malinova and Park (2016)). Market makers and ELPs have an interest in being simultaneously active on both markets. These characteristics allow us to assume that the US price-time prioritization scenario, widely used at the international level, reflects itself in the Canadian price and quote discovery processes.

To evaluate the economic impact of broker preference, we simulate the algorithm activities by applying the price-time prioritization rule in a first step. The broker preference is then applied to Brokers A and B, whose customers have significantly different transaction volumes. We analyze the viability of performances under the two prioritization scenarios and evaluate the influence of the maker/take model in Section 3.5.1. Section 3.5.2 compares the ELP's role as a liquidity provider to the stylized facts. We evaluate the influential market and stock characteristics on ELP's positions in Section 3.5.3. We study the ELP's behavior to determine the illiquidity contagion risk in Section 3.5.4.

3.5.1 Sustainability of performances

At any time, we can determine the exact contents of the LOB including the broker issuing the limit order, its quantity and, its timestamp. This information is also available for all marketable orders. The provisions of Section 3.3.1 and Section 3.3.2 respect the quotes and price discovery processes and we use the same latency for all scenarios. We do not use data mining technics. For these reasons, we are confident that the results of the simulations minimize the gap between our out-of-sample simulations and real-time trading. Figure 3.2 shows the cumulative profits excluding fees and rebates for the price-time scenario and the broker preference rule applied to Brokers A and B. Profits are economically significant and are obtained without high volatility. No significant drawdown is observed.

[insert Figure 3.2 here]

Table 3.4 displays the three simulations aggregated statistics. The price-time prioritization simulation achieves the highest profit at 4.9 m\$ (3.8m\$ without fees and rebates). This is expected as the ELP has a full access to the incoming marketable orders: no privilege is granted to competitors. Broker A obtains the smallest average daily profit of \$25,646. It is nonetheless economically significant. His worst daily loss is \$3,422 even though the maximum one-minute trading loss is \$14,211. This illustrates the algorithm capacity to quickly recover from adverse positions. This characteristic holds for the other two simulations.

[insert Table 3.4 here]

Table 3.5 details the monthly characteristics. The 3.9m trades for A and 2.5m trades for B (Col. 5) reflects the difference in the size of their customers. A provides more liquidity than B and adjusts its quotes more frequently as reflected in the #LO/#trades ratio of 4.89 for A and 3.73 for B. These remain significantly lower than the market ratio of 15.4 (Table 3.3, col. 5). The algorithm conceptualization is at the origin of this result. As we compel the ELP to respect the price and quote discovery processes, he is strictly reacting to other participants induced moves. Moreover, he does not manage orders from other LOB levels as his quoting activities are restricted to the first level.

[insert Table 3.5 here]

Monthly profits are all economically significant. Broker A trading profit is \$ 553k lower than broker B's, despite a 57% higher trading volume. The explanation could be associated with the degree of sophistication of broker A's clients. The ELP faces a more severe adverse selection cost in the sense of Glosten and Milgrom (1985) under broker preference if the broker's clientele have a greater proportion of informed traders. The monthly order-to-trade ratios (Table 3.5, Col. 6) are systematically higher for Broker A than Broker B. This is an indication of a greater presence of HFTers when Broker A modify his liquidity providing positions. Brogaard, Hendershott et al. (2014) note that HFTers trade in the direction of permanent price changes through their liquidity demanding orders, a informed traders characteristic. The maker-taker model mitigates the gap in net profits by leveling the total compensation of both brokers to ≈ 4 million \$ over the six-month period. Broker

A and Broker B brokerage revenues are not available. Considering the large discrepancy in the number of transactions of their clientele, Broker A undoubtedly has a more substantial income. Throughout the stock market correction of August 2015, the algorithm obtains its best performances from all simulations. It is a sign of efficiency from its activities in volatile environments. The price time and broker A performance are more exposed to future declines in profitability as their share of total earnings coming from the maker/taker model (Col. 4) is almost twice that of broker B. Future decreases in the maker/taker fees are already planned by the Toronto Stock Exchange. Concurrently, the brokers experienced their record levels of quotes (Col. 5* Col. 6) and trades (Col. 5). This confirms their presence during high uncertainty periods. Their behaviors are further analyzed in Section 3.5.4.

The parameters of the brokers A and B empirical distributions of P&Ls are also comparable as shown in Table 3.6.

[insert Table 3.6 here]

3.5.2 Stylized facts

To assess the ELPs role as liquidity providers, we refer to a methodology proposed by Malinova, Park et al. (2016) which examines the impact of submission and cancellation fees on market quality. To identify traders who act as market makers, they set a critical threshold to less than 0.20 for the activity index of Equation (3.2).

$$\begin{aligned} & \text{absolute value of market maker index} \\ &= \frac{|passive\ buy\ order\ volume - passive\ sell\ order\ volume|}{|passive\ buy\ order\ volume + passive\ sell\ order\ volume|}, \end{aligned} \quad (3.2)$$

Table 3.7 contains the volume of limit orders from brokers A and B and the market maker index. During the six months of activity, the algorithm maintains a median ratio smaller than 0.20.

[insert Table 3.7 here]

Malinova, Park et al. (2016) note that high-frequency market makers account for 44% of limit orders and 34% of trades. The monthly activities of brokers A and B are compared to actual ones in Table 3.8. Although they represent a high volume of activity, all results are well below the benchmarks.

[insert Table 3.8 here]

3.5.3 Influential market and stocks characteristics

High-frequency market making is of concern to IIROC, the Canadian financial markets regulatory instance (IIROC (2015)). Their request to investigate the liquidity provision by high-frequency market makers resulted in Malinova and Park (2015) paper. The paper confirms the importance of endogenous liquidity providers in Canada. This section focuses to complement the existing (and scarce) literature in the Canadian settings by analyzing the impact of market states on ELP behavior.

The information known at the time the algorithm makes its quoting decisions include the traded market value, the stock idiosyncratic risk, the bid-ask spread, the number of quotes, the imbalance, and the stock price trend. The ELP's positions and their variations are generated by incoming marketable orders and level 1 market states observed. Recent empirical evidence shows that a number of high-frequency variables are affected by persistence: momentum and contrarian regimes (Andersen, Cebiroglu et al. (2017)), order flow (Tóth, Palit et al. (2015)), and volatilities (Patton and Sheppard (2015); Gatheral, Jaisson et al. (2016)). Salighehdar, Liu et al. (2017) find clusters in high-frequency liquidity measures and Yang and Zhu (2016) in limit orders. We use first differences with market values, bid-ask spreads, limit orders, and imbalances to get weakly dependent processes integrated of order zero. We apply a percentage of change $((x_t - x_{t-1})/x_{t-1})$ to standard deviations and momentum measures (Wooldridge (2005), p.406).

We assess the market states influencing the variation in ELP's position for stock i , on day t , from time interval $t - 1$ to time interval t ($diffMVelp_{i,d,t}$) using several specifications of Equation (3.3) applied to brokers A and B.

$$\begin{aligned} diffMVelp_{i,d,t} = & \beta_1 diffMV_{i,d,t} + \beta_2 pct\sigma_{i,d,t} + \beta_3 diffSprd_{i,d,t} + \\ & \beta_4 diff\#Quotes_{i,d,t} + \beta_5 diffImbal_{i,d,t} + \beta_6 pctMom_{m,d,t} + \varepsilon_{i,d,t}. \end{aligned} \quad (3.3)$$

On the right-hand side $diffMV_{i,d,t}$ is the first difference in total market value traded for stock i on day d from time interval $t - 1$ to time interval t , a measure of trade intensity. $pct\sigma_{i,d,t}$ is the percentage of change of stock i standard deviation on day d from time interval $t - 1$ to time interval t , a measure of relative variation in idiosyncratic risk. $diffSprd_{i,d,t}$ is the first difference in bid-ask spread for stock i on day d from time interval $t-1$ to time interval t , a measure of ELP remuneration. $diffImbal_{i,d,t}$ is the first difference in LOB level 1 imbalance on day d from time interval $t - 1$ to time interval t , a liquidity imbalance measure. $diff\#Quotes_{i,d,t}$ is the first difference in the number of quotes for stock i on day d from time interval $t - 1$ to time interval t , a measure of liquidity competition. $pctMom_{m,d,t}$ is the percentage of change in S&P/TSX 60 index momentum on day d from time interval t to time interval $t - 1$. Momentum is defined as $abs\left((up\ vol_{m,d,t}/tot\ vol_{m,d,t}) - 0.50\right)$. It indicates the S&P/TSX 60 index directional intensity. We apply Cameron, Gelbach et al. (2011) procedure which cluster the standard errors at the firm and time interval to account for time-series and cross-sectional dependence. Matlab codes are from Gow, Ormazabal et al. (2010).

Table 3.9 presents the results of multiple specifications from Equation (3.3) for Brokers A and B. As expected, trading intensity measured by $diffMV_{i,d,t}$, is positively correlated to the ELP position variations. This relation is statistically significant for all regression specifications and brokers. We find a linear, positive and significant relationship between changes in equity price volatility and ELP positions. Increases in volatility create an incentive to participation, a sign of ELP's presence in volatile environments for all model specifications and both brokers. ELP positions are positively correlated to the variations in liquidity premiums ($diffSprd_{i,d,t}$). Increases in remuneration appear sufficient to compensate for illiquidity risk. Coefficients of $diff\#Quotes_{i,d,t}$ are positive and statistically significant. With the Ait-Sahalia and Saglam (2014) model, the algorithm

increases (decreases) the optimal quoting quantities according to the increase (decrease) in the number of quote updates, a main component of ELP signals. Temporary liquidity disequilibrium does not affect broker positions as variations in imbalances ($diffImbal_{i,d,t}$) are not statistically significant. Acknowledging statistically non-significant loadings, the algorithm exhibits a weak contrarian strategy as ELP positions are negatively correlated to $pctMom_{m,d,t}$. All Equation () idiosyncratic variables are taken into consideration while determining the OQPs, except volatility. An extension to Ait-Sahalia and Saglam (2014) model including this variable should improve its performance.

[insert Table 3.9 here]

3.5.4 Illiquidity contagion risk

The aim of this section is to study ELP's behavior to determine the risk of withdrawal under stressful market conditions. In order to minimize market risk, the ELPs target zero inventory. Liquidity providers are under pressure when stock prices are strongly directional and trading intensity high. These situations limit bid-ask bounces, and result in high inventory levels and market risk for the liquidity providers. That could prompt ELPs to temporarily cease their activities, increasing the probability of illiquidity contagion.

Market states are monitored with two characteristics: the number of trades and the absolute logarithmic return. Critical situations are those where the value of these characteristics is greater than their quantile values at 90%. 38 272 one-minute time intervals have one or more stocks with high trend intensity and directional movement ($\approx 1.4\%$ of the total). The risk of contagion is exacerbated when several stocks meet these two criteria simultaneously. In order to focus on extreme events, the critical threshold we use is fifteen stocks, i.e. at least 25% of the stocks from our sample must exceed simultaneously their 90% quantile values.

[insert Figure 3.3 here]

Figure 3.3 visualizes the distribution of the 572 intervals having between fifteen and forty-nine stocks exceeding the 90% thresholds. 92.8% of occurrences involve between

15 and 30 stocks simultaneously. Unsurprisingly, the bulk of critical market states occurs in August as depicted by Figure 3.4.

[insert Figure 3.4 here]

Liquidity providers learn information about a stock from the behavior of other stocks. According to Cespa and Foucault (2014), this generates self-reinforced feedback loops that cause illiquidity contagion. Our measure of liquidity is the average quantity displayed on a given side at level one. Since we do not want to influence our liquidity measures due to trend in prices, we use quantities rather than market values. Figure 3.5 presents market statistics for the ask by one-minute time intervals during April 2015. The minimum quantity occurs at the opening of the session. It reaches a peak after about 45 minutes, drops to 10,000 units and remains there until 30 minutes before the close, while supply increases significantly. The bid statistics of August 2015 depicted in Figure 3.6 exhibit the same behavior.

[insert Figure 3.5 here]

[insert Figure 3.6 here]

To take into account the presence of liquidity cyclicity, we use Equation (3.4) as a measure of the relative ELP participation to liquidity:

$$ELPbid_{int,d,i} = \frac{bid_{int,d,i}}{mean(bid_{int,i})}, \quad (3.4)$$

where:

$ELPbid_{int,d,i}$: Relative ELP participation to bid liquidity over interval int on day d for stock i .

$bid_{int,d,i}$: ELP's quantity on the bid over interval int on day d for stock i .

$mean(bid_{int,i})$: ELP's average quantity on the bid over interval int for stock i .

The relative ELP participation to the ask is calculated the same way. Table 3.10 presents the cross-correlation matrix of the relative ELP participation to liquidity when he is acting as broker B during the 575 extreme situations identified. The upper part of the diagonal

matrix displays the cross-correlation coefficients of the relative bid quantities while the lower part contains the cross-correlation coefficients of the relative offer quantities. The last row and the last column include their respective averages.

[insert Table 3.10 here]

Values of the cross-correlation coefficients range from -0.1090 to 0.5522 and 97.8% are positive. We note that the cross-correlation matrix distribution is centered around a value slightly below +0.20, which implies positively correlated behavior of ELP's liquidity from both sides of the market (Figure 3.7).

[insert Figure 3.7here]

There are no significant differences between bid and ask. Detailed analysis of critical situations reveals eight time intervals where the ELP has completely stopped providing liquidity. Of these, seven occur in the last two minutes of trading in continuous time. They are due to risk management features to close positions before the end of the day. The last one occurred on August 25, 2015, at 10:27. It cannot be dissociated from the fundamental information content that caused a starting market downturn during the previous day. Quoting activities are back to normal as of the following period.

3.6. Conclusion

With an algorithm conceptualized by Yergeau (2016), we evaluate the impact of broker preference on the performance of endogenous liquidity providers. Using the international standard of price-time hierarchy as a benchmark and the characteristics of two brokers with a significantly different clientele, we establish the economic viability of ELPs. Profitability is high, even excluding rebates to liquidity providers. P&Ls volatility is low and no significant drawdown occurs. While comparing with the price-time benchmark, Brokers A and B obtain smaller revenues including rebates, fees, and commissions under the Canadian broker preference feature. An explanation to this result is the ELPs unconstrained access to all marketable orders under price-time priority. This prioritization induces the speed arms race invoked by Budish, Cramton et al. (2015), Schwartz and Wu (2013), and Wah, Hurd et al. (2015). As broker preference gives priority over other

brokers' limit orders, incentive to invest in latency reduction is strongly mitigated, thus reducing the gap between the two scenarios. Analyzing performance under broker preference, gross trading profits of Broker B are higher than Broker A, even though Broker B have a smaller clientele. This is inversed after considering fees and rebates from the maker/taker model and broker commissions. Analysis of the monthly statistics shows that the best performance arises in August during a stock market correction, a sign of the algorithm efficiency in volatile environment. August is also the month when ELPs provide the most liquidity. This mitigates the worries of evasive liquidity.

The analysis of stylized facts shows that the algorithm exhibits the characteristics of endogenous liquidity providers. The ELP positions are positively correlated with the intensity of trades and quotes, and the bid-ask spread. His activities are also positively correlated with the idiosyncratic risk of stocks, a hint to ELP presence during volatile markets. Liquidity imbalances and market momentum do not significantly influence their positions. We analyze the behavior of ELPs in extreme situations. The risk of illiquidity contagion is quantified using a cross-correlation matrix. The coefficients are centered slightly below +0.20. Although a detailed analysis of the critical events shows that ELPs did not withdraw from the market during the period covered by our sample, we cannot exclude such a situation in the future.

References

- Ait-Sahalia, Yacine, and Mehmet Saglam. 2014. "High Frequency Traders: Taking Advantage of Speed." *SSRN Electronic Journal*.
- Anand, Amber, and Kumar Venkataraman. 2016. "Market Conditions, Fragility, and the Economics of Market Making." *Journal of Financial Economics* no. 121 (2):327-349.
- Andersen, Torben G., Gökhan Cebiroglu, and Nikolaus Hautsch. 2017. "Volatility, Information Feedback and Market Microstructure Noise: A Tale of Two Regimes." *Center for Financial Studies, Goethe University (Working Paper No. 569)*.
- Baldauf, Markus, and Joshua Mollner. 2016. "Fast Traders Make a Quick Buck: The Role of Speed in Liquidity Provision." *SSRN Electronic Journal*.
- Basel Committee on Banking Supervision. 2013. "Fundamental Review of the Trading Book: A Revised Market Risk Framework." *Research Report*.
- Black Rock. 2015. US Equity Market Structure: Lessons from August 24th. *Research Report*.
- Bouchaud, J.P., J.D. Farmer, and F. Lillo. 2009. "How Markets Slowly Digest Changes in Supply and Demand." In *Handbook of Financial Markets: Dynamics and Evolution*, edited by T. Hens and K.R. Schenk-Hoppé, pp. 57–160.
- Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan. 2014. "High Frequency Trading and Price Discovery." *Review of Financial Studies* no. 27 (8):2267-2306.
- Budish, Eric, Peter Cramton, and John Shim. 2015. "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response." *The Quarterly Journal of Economics* no. 130 (4):1547-1621.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business & Economic Statistics* no. 29 (2):238-249.

Cespa, Giovanni, and Thierry Foucault. 2014. "Illiquidity Contagion and Liquidity Crashes." *The Review of Financial Studies* no. 27 (6):1615-1660.

Chordia, Tarun, Amit Goyal, Bruce N. Lehmann, and Gideon Saar. 2013. "High-frequency Trading." *Journal of Financial Markets* no. 16 (4):637-645.

Chung, Kee H., and Albert J. Lee. 2016. "High-frequency Trading: Review of the Literature and Regulatory Initiatives around the World." *Asia-Pacific Journal of Financial Studies* no. 45 (1):7-33.

Cont, Rama , Arseniy Kukanov, and Sasha Stoikov. 2014. "The Price Impact of Order Book Events." *Journal of Financial Econometrics* no. 12 (1):47-88.

Frijns, Bart, Ivan Indriawan, and Alireza Tourani-Rad. 2014. "Macroeconomic News Announcements and Price Discovery: Evidence from Canadian–U.S. Cross-listed Firms." *Journal of Empirical Finance* no. 32:35-48.

Gatheral, Jim, Thibault Jaisson, and Mathieu Rosenbaum. 2016. "Volatility is Rough." *Cornell University Library - Statistical Finance*.

Glosten, Lawrence R., and Paul R. Milgrom. 1985. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics* no. 14 (1):71-100.

Gow, Ian D., Gaizka Ormazabal, and Daniel J. Taylor. 2010. "Correcting for Cross-Sectional and Time-series Dependence in Accounting Research." *The Accounting Review* no. 85 (2):483-512.

Hautsch, Nikolaus, and Ruihong Huang. 2012. "The Market Impact of a Limit Order." *Journal of Economic Dynamics & Control* no. 36:501-522.

Hope, B. 2013. "Upstairs' Market Draws More Big Investors." *Wall Street Journal (Online)*, December 7 (<http://online.wsj.com/news/articles/SB10001424052702303722104579240393018918698>).

Huh, Sahn-Wook. 2014. "Price Impact and Asset Pricing." *Journal of Financial Markets* no. 19:1-38.

IIROC. 2014a. "Fee Model Guidelines - Update." *Notice 14-0284*.

IIROC. 2014b. "Identifying Trading Groups – Methodology and Results." *IIROC Notice 14-0210*.

IIROC. 2015. "IIROC's Study of High Frequency Trading (HFT)." *IIROC Forum on High Frequency Trading (HFT) October 19*.

IIROC. 2016. "Statistics on Canadian Stock Markets." Available at: <http://www.iroc.ca/industry/marketmonitoringanalysis/Pages/StatisticsInformation.aspx>

Irwin, Neil. 2015. "Why the Stock Market Is so Turbulent." *The New York Times* (August 24) (Retrieved 2017/02/01 from: <https://www.nytimes.com/2015/08/25/upshot/why-global-financial-markets-are-going-crazy.html?ribbon-ad-idx=3&rref=business&module=ArrowsNav&contentCollection=DealBook&action=swipe®ion=FixedRight&pgtype=article&abt=0002&abg=1&r=0>).

Jain, Pankaj K., Pawan Jain, and Thomas H. McInish. 2016. "Does High-frequency Trading Increase Systemic Risk?" *Journal of Financial Markets* no. 31:1-24.

Korajczyk, Robert A., and Dermot Murphy. 2016. "High Frequency Market Making to Large Institutional Trades." *SSRN Electronic Journal*.

Lopez De Prado, Marcos. 2015. "The Future of Empirical Finance." *Journal of Portfolio Management* no. 41 (4):140-144.

Malinova, Katya, and Andreas Park. 2015. "Liquidity Provision and Market Making by HFTs." *Report prepared for the Investment Industry Regulatory Organization of Canada*.

Malinova, Katya, and Andreas Park. 2016. "'Modern' Market Makers." *University of Toronto. Working Paper*.

Malinova, Katya, Andreas Park, and Ryan Riordan. 2014. "Do Retail Traders Suffer from High Frequency Traders?" *SSRN Electronic Journal*.

Malinova, Katya, Andreas Park, and Ryan Riordan. 2016. "Taxing High Frequency Market Making: Who Pays the Bill?" *SSRN Electronic Journal*.

Menkveld, Albert J. 2013. "High Frequency Trading and the New Market Makers." *Journal of Financial Markets* no. 16 (4):712-740.

Patton, Andrew J., and Kevin Sheppard. 2015. "Good Volatility, Bad Volatility: Signed Jumps and Persistence of Volatility." *Review of Economics & Statistics* no. 97 (3):683-697.

S&P Indices. 2010. "A VIX for Canada." *S&P Presentation on October 14, 2010*.

Salighehdar, Amin, Yang Liu, Dragos Bozdog, and Ionut Florescu. 2017. "Cluster Analysis of Liquidity Measures in Stock Market." *Journal of Management Science and Business Intelligence* no. 2 (2):1-8.

Schwartz, Robert A., and Liuren Wu. 2013. "Equity Trading in the Fast Lane: The Staccato Alternative." *Journal of Portfolio Management* no. 39 (3):3-6.

Shan Li, Andrea Chang, and Paresh Dave. 2015. "Stock Market Suffers Worst One-day Drop since 2008." *Los Angeles Times* (August 21) (Retrieved 2017/02/01 from: <http://www.latimes.com/business/la-fi-0822-financial-markets-20150821-story.html>).

Sofianos, G., and J. Xiang. 2013. "Do Algo Executions Leak Information?" In *High-Frequency Trading*. Edited by: Easley, David, Lopez de Prado, Marcos and O'Hara, Maureen. Risk Books.

Swedroe, Larry. 2016. High Frequency Trading's Impact. In *The BAM Alliance* 03/17/2016. <http://thebamalliance.com/blog/high-frequency-tradings-impact/>.

The Government Office for Science London. 2012. "EIA4: Circuit Breakers." *Research Paper*.

TMX Group. 2015a. "Order Types and Functionality Guide - TMX Equity Markets " *June 2015*.

TMX Group. 2015b. "Taking Action to Optimize Maker-taker Fees." *May 2015*.

Tóth, Bence, Imon Palit, Fabrizio Lillo, and J. Doyne Farmer. 2015. "Why Is Equity Order Flow so Persistent?" *Journal of Economic Dynamics and Control* no. 51:218-239.

U.S. Securities and Exchange Commission. 2010. "Concept Release on Equity Market Structure." *Release No. 34-61358* File No. 57-02-10.

Wah, E., D. R. Hurd, and M. P. Wellman. 2015. "Strategic Market Choice: Frequent Call Markets vs. Continuous Double Auctions for Fast and Slow Traders." *In Proceedings of the Conference on Auctions, Market Mechanisms and Their Applications (AMMA)*.

Wooldridge, J. M. 2005. "Introductory Econometrics: A New Approach." Ed. *South-Western Cengage Learning*. 888 pages

Yang, Liyan, and Haoxiang Zhu. 2016. "Back-Running: Seeking and Hiding Fundamental Information in Order Flows." *Rothman School of Management Working Paper No. 2583915*.

Yergeau, Gabriel. 2016. "Profitability and Market Quality of High Frequency Market-Makers: An Empirical Investigation." *SSRN Electronic Journal*.

Zhou, Wei-Xing. 2012. "Universal Price Impact Functions of Individual Trades in an Order-driven Market." *Quantitative Finance* no. 12 (8):1253-1263.

Appendix 3.1 Optimal quoting policy¹¹

Ait-Sahalia and Saglam (2014) modelize two types of agent. Low-frequency traders (LFTers) use exclusively market orders. One ELP exhibits inventory aversion and focuses specifically on liquidity supply. Decisions to provide liquidity depend on the trade-off between the inflows from the bid-ask spread and the outflows associated with inventory. Equation (A3.1.1) describes the expected reward function:

$$E(\pi) = C/2 \sum_{t=1}^{\infty} e^{-Dsmo_t} I(l_{smo_t}^b = 1) + C/2 \sum_{t=1}^{\infty} e^{-Dbmo_t} I(l_{bmo_t}^a = 1) - \Gamma \int_0^{\infty} e^{-Dt} |x_t| dt, \quad (A3.1.1)$$

where:

$E(\pi)$: Quoting policy expected reward.

C : Bid-ask spread.

D : Constant discount factor > 0 .

smo_t (bmo_t): Sell (buy) market order by LFTs at time t .

I : Indicator function.

b : ELP bid limit order.

a : ELP offer limit order.

$l_{smo_t|bmo_t}^{b|a}$: Equals 1 if the ELP is quoting a bid (b) or an offer (a) limit order when a LFT sell (buy) market order arrives, 0 otherwise.

Γ : Inventory aversion coefficient.

x_t : Inventory position at time t .

The first term to the right of Equation (A3.1.1) is the discounted value of the ELP's revenue $\left(\frac{C}{2}\right)$ earned when an incoming LFT's sell market order hits the ELP's limit order while he is bidding $\left(I(l_{smo_t}^b = 1)\right)$. The second term is the discounted revenue associated with an incoming LFT's bid market order, and the third term is the discounted value of the ELP's inventory costs over the period dt . To keep the model tractable, the ELP always places his LOs at the best bid and/or offer price and does not issue orders larger than one contract.

¹¹ This Appendix is based on Yergeau, G. (2016). "Profitability and Market Quality of High Frequency Market-Makers: An Empirical Investigation." SSRN Electronic Journal, Section 2.

Apart from observing the arrival of market orders, the ELP receives a signal s about the likely side of the next incoming market order: $s \in \{1, -1\}$, where 1 predicts an incoming LFT's buy market order and -1 an incoming LFT's sell market order. P quantifies the informational quality of the ELP's signal. It varies from 0.5 (no prior knowledge about the side of the next incoming LFT's market order) to 1.0 (perfect knowledge). In Ait-Sahalia and Saglam (2014) setup, the next event is either 1: the arrival of a signal with probability $\left(\frac{\mu/2}{\lambda+\mu}\right)$, μ being the arrival rate of a Poisson distribution of the ELP's signals and λ the arrival rate of a Poisson distribution of the incoming LFTs' market orders; 2: the arrival of a market order in the direction of the last signal with probability $\frac{P\lambda}{\lambda+\mu}$; or 3: the arrival of a market order in the opposite direction of the last signal with probability $\frac{(1-P)\lambda}{\lambda+\mu}$. The value of market-making activities for any given event assuming an inventory position of x ($x \in \{\dots, -2, -1, 0, 1, 2, \dots\}$) and a sell signal (-1) is:

$$v(x, -1) = -\gamma|x| + \delta \left\{ \left(\frac{\mu/2}{\lambda+\mu} \right) (v(x, 1) + v(x, -1)) + \frac{P\lambda}{\lambda+\mu} \max \left(\frac{c}{2\delta} + v(x-1, -1), v(x, -1) \right) + \frac{(1-P)\lambda}{\lambda+\mu} \max \left(\frac{c}{2\delta} + v(x+1, -1), v(x, -1) \right) \right\}, \quad (\text{A3.1.2})$$

where:

$$\gamma = \frac{\Gamma}{\lambda+\mu+D}; \quad \delta = \frac{\lambda+\mu}{\lambda+\mu+D}; \quad c = \delta C.$$

Equation (A3.1.2) quantifies the market-making value function. The first term to the right is the discounted inventory cost ($-\gamma|x|$). The second term is the discounted value of the three possible events: the value of the arrival of a signal $\left(\left(\frac{\mu/2}{\lambda+\mu} \right) (v(x, 1) + v(x, -1)) \right)$, the value of the arrival of a market order in the direction of the signal $\left(\frac{P\lambda}{\lambda+\mu} \max \left(\frac{c}{2\delta} + v(x-1, -1), v(x, -1) \right) \right)$, and the value of the arrival of a market order in the opposite direction of the signal $\left(\frac{(1-P)\lambda}{\lambda+\mu} \max \left(\frac{c}{2\delta} + v(x+1, -1), v(x, -1) \right) \right)$.

Solving Equation (A3.1.2) by backward induction using the Hamilton-Jacobi-Belman optimality method leads to the optimization of the expected reward trade-off.

According to Ait-Sahalia and Saglam (2014)'s *Theorem 1*, the OQP consists in quoting the best bid and/or ask price(s) according to a threshold policy:

$$\begin{aligned} l^b(x, 1) &= \begin{cases} 1 & \text{when } x < U^* \\ 0 & \text{when } x \geq U^* \end{cases} & l^a(x, 1) &= \begin{cases} 1 & \text{when } x > L^* \\ 0 & \text{when } x \leq L^* \end{cases} \\ l^b(x, -1) &= \begin{cases} 1 & \text{when } x < -L^* \\ 0 & \text{when } x \geq -L^* \end{cases} & l^a(x, -1) &= \begin{cases} 1 & \text{when } x > -U^* \\ 0 & \text{when } x \leq -U^* \end{cases} \end{aligned}$$

where U^*, L^* are respectively the bid and offer optimal quantities.

Theorem 1 can be interpreted as follows: Suppose the ELP receives a “buy” signal ($s = 1$) while being long ($x > 1$). He is going to act upon it ($l^b = 1$) as long as its current inventory is not already too high ($x < U^*$). If ($x \geq U^*$), the ELP will not quote because this could increase its long inventory position beyond the optimal threshold U^* . Symmetrically, if the ELP receives a “sell” signal ($s = -1$), he will quote on the offer side ($l^a = 1$) as long as its inventory position is not already too short ($x > -U^*$).

An algorithm proposed by Ait-Sahalia and Saglam (2014) presented in their Appendix allows us to determine the thresholds based on the expected reward trade-off and *Theorem 1*.

Appendix 3.2 Observed parameters value

A3.2. 1 Trades by minute first 5 days, March 2015

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	avg.	std. dev.	median	min.	max.	95%
Ticker	#	#	#	#	#	quantile
ABX	15	18	11	-	282	81
AEM	11	14	7	-	247	60
AGU	6	7	4	-	122	33
ARX	13	13	10	-	162	58
ATD.B	14	15	10	-	199	69
BAM.A	11	13	8	-	166	63
BB	14	18	9	-	359	85
BBD.B	12	25	3	-	481	105
BCE	18	18	13	-	270	86
BMO	20	19	16	-	243	92
BNS	25	24	18	-	427	119
CCO	9	11	5	-	137	52
CM	15	16	11	-	234	75
CNQ	26	25	19	-	359	119
CNR	17	16	13	-	248	78
CP	6	8	4	-	101	35
CPG	22	22	16	-	353	99
CSU	1	2	-	-	32	10
CTC.A	3	5	1	-	63	24
CVE	21	21	15	-	285	103
DOL	4	6	1	-	88	28
ECA	18	22	12	-	503	100
ELD	11	16	6	-	473	69
EMA	5	8	2	-	110	36
ENB	19	19	14	-	284	89
FM	20	22	14	-	275	102
FNV	7	8	5	-	129	39
FTS	11	12	7	-	157	55
G	20	22	14	-	356	101
GIB.A	10	11	7	-	149	52
GIL	5	6	2	-	75	30
HSE	12	13	8	-	183	59
IMO	11	12	8	-	171	56
IPL	10	11	7	-	127	50
K	7	13	1	-	147	61
L	8	10	6	-	104	44
MFC	16	17	11	-	244	78

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	avg.	std. dev.	median	min.	max.	95%
Ticker	#		#	#	#	quantile
MG	7	10	5	-	329	41
MRU	8	10	5	-	169	42
NA	13	14	9	-	198	65
POW	9	11	6	-	184	50
PPL	13	13	10	-	222	58
QSR	6	8	3	-	103	38
RCI.B	11	12	8	-	141	56
RY	26	25	19	-	412	122
SAP	5	8	2	-	104	38
SJR.B	8	12	5	-	196	57
SLF	14	16	10	-	269	71
SLW	16	17	11	-	239	80
SNC	9	12	5	-	217	49
SU	27	26	19	-	451	127
T	14	16	10	-	270	78
TCK.B	20	31	13	-	1,103	121
TD	26	26	19	-	318	125
TRI	10	11	6	-	143	55
TRP	14	15	10	-	191	71
VRX	7	9	4	-	170	42
WN	2	4	-	-	63	19
YRI	9	15	3	-	176	71

All statistics refer to 1-minute time intervals; Avg. # : average number of trades; std. Dev.: standard deviation of the number of trades; median # : median number of trades; min. # : minimum number of trades; max. # : maximum number of trades; 95% quantile : 95% quantile value of the number of trades.

A3.2. 2 Signals by minute, first 5 days, March 2015

(1)	(2)	(3)	(4)	(5)	(6)	(7)
	avg.	std. dev.	median	min.	max.	95%
Ticker	#	#	#	#	#	quantile
ABX	297	644	250	13	50,013	1,172
AEM	160	86	145	35	2,586	470
AGU	234	55	240	-	1,136	384
ARX	149	89	145	-	3,901	398
ATD.B	175	121	161	-	8,859	448
BAM.A	179	76	169	66	5,179	343
BB	311	320	264	11	10,013	1,277
BBD.B	778	1,380	319	2	41,844	5,010
BCE	205	134	188	53	8,579	469
BMO	211	55	199	75	1,148	409
BNS	214	79	199	62	2,883	463
CCO	162	110	152	17	1,756	519
CM	223	66	206	91	1,193	473
CNQ	220	165	198	36	9,672	563
CNR	216	62	203	83	1,574	434
CP	328	56	336	-	1,044	476
CPG	192	106	170	27	2,214	560
CSU	370	200	437	-	2,038	634
CTC.A	193	66	228	-	1,129	338
CVE	202	333	179	21	27,482	592
DOL	144	129	163	-	6,865	505
ECA	280	269	230	14	5,340	1,176
ELD	329	617	223	-	26,373	1,921
EMA	118	72	141	-	1,216	316
ENB	186	67	176	57	3,636	361
FM	195	127	186	-	3,052	571
FNV	161	66	163	-	1,671	360
FTS	151	131	148	-	10,099	340
G	226	153	196	23	5,607	694
GIB.A	165	70	155	-	1,554	379
GIL	154	67	175	-	1,775	327
HSE	152	90	142	-	2,816	428
IMO	162	66	155	-	2,071	350
IPL	154	101	145	-	4,153	467
K	318	649	103	-	25,904	2,795
L	170	88	164	-	4,063	392
MFC	290	228	247	21	5,021	1,141

RCI.B	159	119	150	42	9,428	359
RY	231	106	212	75	3,932	488
SAP	122	93	135	-	1,553	412
SJR.B	147	168	145	-	12,946	384
SLF	184	94	172	38	3,976	440
SLW	183	85	170	23	2,127	456
SNC	147	109	141	-	4,015	459
SU	222	87	204	-	1,673	537
T	177	111	161	41	4,829	464
TCK.B	223	178	203	17	7,817	694
TD	237	124	214	53	5,833	564
TRI	169	83	165	49	3,829	368
TRP	183	139	170	53	9,817	396
VRX	353	57	355	-	1,747	507
WN	139	65	104	-	1,177	301
YRI	312	436	188	4	6,393	2,005

All statistics refer to 1-minute time intervals; Avg. # : average number of signals; std. Dev.: standard deviation of the number of signals; median # : median number of signals; min. # : minimum number of signals; max. # : maximum number of signals; 95% quantile : 95% quantile value of the number of signals.

A3.2. 3 Spreads by minute first 5 days, March 2015

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ticker	# spreads	avg. price	avg. q	avg. spread	std. dev. spread	median spread	95% quantile
ABX	722,340	15.383	3,421	0.0109	0.0058	0.0100	0.0200
AEM	435,530	38.770	299	0.0255	0.0220	0.0200	0.0700
AGU	256,611	141.770	140	0.1001	0.0851	0.0800	0.4000
ARX	225,848	23.967	429	0.0174	0.0191	0.0100	0.0600
ATD.B	200,087	47.407	259	0.0301	0.0243	0.0300	0.1000
BAM.A	291,756	68.264	263	0.0332	0.0321	0.0300	0.1100
BB	709,952	13.721	3,782	0.0114	0.0051	0.0100	0.0200
BBD.B	177,388	2.394	72,591	0.0105	0.0027	0.0100	0.0200
BCE	823,474	54.826	729	0.0140	0.0099	0.0100	0.0300
BMO	672,780	76.345	408	0.0200	0.0157	0.0200	0.0600
BNS	670,748	65.442	540	0.0158	0.0166	0.0100	0.0400
CCO	369,196	18.760	1,035	0.0125	0.0086	0.0100	0.0300
CM	347,743	94.480	252	0.0339	0.0253	0.0300	0.1100
CNQ	871,422	37.234	770	0.0148	0.0129	0.0100	0.0400
CNR	443,546	85.971	267	0.0297	0.0277	0.0300	0.1000
CP	308,822	236.901	124	0.2124	0.1706	0.1800	0.7600
CPG	490,697	30.066	591	0.0140	0.0097	0.0100	0.0400
CSU	49,977	429.620	114	2.1045	1.2625	1.9000	6.6600
CTC.A	89,941	129.621	136	0.1911	0.1192	0.1800	0.5600
CVE	593,058	21.898	1,819	0.0116	0.0065	0.0100	0.0200
DOL	50,195	62.966	155	0.0638	0.0495	0.0600	0.2600
ECA	470,251	15.303	3,920	0.0113	0.0082	0.0100	0.0200
ELD	366,051	6.478	5,663	0.0109	0.0057	0.0100	0.0200
EMA	119,025	41.342	256	0.0310	0.0299	0.0300	0.1100
ENB	424,323	58.151	336	0.0218	0.0202	0.0200	0.0800
FM	384,943	15.503	1,122	0.0131	0.0145	0.0100	0.0400
FNV	210,013	63.923	157	0.0534	0.0572	0.0400	0.2000
FTS	227,001	39.183	440	0.0183	0.0183	0.0100	0.0600
G	983,751	25.703	1,363	0.0123	0.0099	0.0100	0.0200
GIB.A	258,471	53.082	255	0.0353	0.0377	0.0300	0.1200
GIL	159,675	75.017	180	0.0570	0.0481	0.0500	0.2300
HSE	292,722	27.617	495	0.0151	0.0129	0.0100	0.0500
IMO	183,889	47.851	221	0.0311	0.0337	0.0300	0.1000
IPL	155,419	33.107	370	0.0205	0.0183	0.0200	0.0800
K	165,979	3.387	25,579	0.0103	0.0044	0.0100	0.0200
L	128,899	62.279	166	0.0507	0.0464	0.0500	0.1600
MRU	108,904	34.918	256	0.0294	0.0289	0.0200	0.1100
NA	306,859	47.759	350	0.0205	0.0162	0.0200	0.0700
POT	828,646	43.772	1,003	0.0132	0.0101	0.0100	0.0400
POW	262,498	33.415	456	0.0156	0.0132	0.0100	0.0500

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ticker	# spreads	avg. price	avg. q	avg. spread	std. dev. spread	median spread	95% quantile
PPL	314,884	41.260	320	0.0216	0.0194	0.0200	0.0600
QSR	130,894	53.571	176	0.0559	0.0675	0.0400	0.2600
RCI.B	274,561	43.782	376	0.0173	0.0157	0.0200	0.0600
RY	951,518	77.353	503	0.0170	0.0120	0.0200	0.0400
SAP	122,412	36.370	276	0.0274	0.0233	0.0200	0.0900
SJR.B	265,632	28.904	926	0.0132	0.0115	0.0100	0.0400
SLF	479,332	38.870	809	0.0141	0.0120	0.0100	0.0400
SLW	736,470	24.906	1,006	0.0128	0.0101	0.0100	0.0300
SNC	121,149	38.308	242	0.0380	0.0472	0.0300	0.1500
SU	879,353	37.174	1,053	0.0129	0.0093	0.0100	0.0300
T	381,194	43.813	397	0.0172	0.0129	0.0200	0.0500
TCK.B	639,717	19.105	1,798	0.0119	0.0096	0.0100	0.0300
TD	1,196,194	54.290	943	0.0132	0.0078	0.0100	0.0300
TRI	353,478	49.415	514	0.0174	0.0318	0.0100	0.0600
TRP	329,070	55.213	394	0.0206	0.0256	0.0200	0.0700
VRX	509,542	253.025	126	0.1868	0.1559	0.1600	0.6400
WN	125,027	103.903	134	0.2016	0.1584	0.1900	0.5800
YRI	183,665	5.115	9,102	0.0107	0.0048	0.0100	0.0200

All statistics refer to 1-minute time intervals and level 1 quotes; # spreads: total number of spreads; avg. price: average $((bid+ask)/2)$; avg. q: average (minimum bid-ask quantity; avg. spread: average value of the spread; std. dev. spread: standard deviation of the spread; median spread : median value of the spread; 95% quantile : 95% quantile value of the spreads.

Table 3. 1 Sessions on the TSX

Time	Session
7:00 AM – 9:30 AM	Pre-Open – Orders may be entered, but will not be executed. The calculated opening price is displayed and continuously updated.
9:30 AM	Market On Open (MOO) – All tradable orders are executed at a single opening trade price with any remaining orders to carry through to the continuous trading session.
9:30 AM – 4:00 PM	Continuous trading – All regular order types are accepted.
3:40 PM – 4:00 PM	Market On Close – MOC Market and Limit on Close (LOC) orders are accepted without restriction until 3:40 PM. The MOC imbalance is published at 3:40 PM, after which only LOC orders opposite to the imbalance side are accepted. Accepted trades will be executed and published at 4:00 PM unless a Price Movement Extension is required when additional offsetting liquidity is solicited, in which case trades are published at 4:10 PM.
4:10 PM – 4:15 PM	Post Market Cancellation Session – During this session, open orders may be canceled by the dealer.
4:15 PM – 5:00 PM	Extended Trading Session – Orders at the last sale price are accepted, but the trades may only occur at the last sale price. Day orders participate in this session. The Must Be Filled (MBF) session for option expiry takes place during Extended Trading once per month, in the evening before an option expiry day.

Ref.:TMX Group (2015), p.8.

Table 3. 2 TSX: maker-taker rebate and fee levels

	Before June 1,2015		After June 1,2015	
	Taker fee	Maker rebate	Taker fee	Maker rebate
Equities priced ≥ 1\$ (per share traded)	\$0.0035	\$0.0031	\$0.0023	\$0.0019

Ref.: TMX Group (2015), p. 3.

Table 3. 3 Trades and quotes : March - August 2015

	(1) Trades M.V. (,000\$)	(2) M.V. (%)	(3) Trades (,000)	(4) Quotes (,000)	(5) Quotes/ Trades	(6) Trades Avg. Q.	(7) Level 1 Avg. Q.	(8) Avg. # orders
March	47,345,942	18.5%	6,244	107,670	17.2	216	3,526	7
April	44,183,491	17.3%	5,570	99,615	17.9	222	3,799	8
May	28,227,624	11.0%	3,682	61,021	16.6	211	3,473	8
June	43,097,374	16.8%	5,744	82,006	14.3	203	3,588	7
July	42,274,678	16.5%	6,118	84,090	13.7	226	3,400	7
August	50,691,069	19.8%	7,687	106,356	13.8	226	3,396	7
	255,820,179	100.0%	35,045	540,758	15.4	217	3,530	7

This table supplies a summary of market statistics. All quotes statistics are based on LOB level 1 only; M.V. stands for market values; column (1) is the total market value traded; column (2) is the percentage of the total market value traded for a given month over the total market value traded over the full sample; column (3) is the number of trades; column (4) is the number of quotes; column (5) is column (4) over column (3); column (6) is the trade average quantity; column (7) is the average level 1 quantity; column (8) is the average number of orders at LOB's level 1.

Table 3. 4 Global performance, price-time priority vs broker preference March-August 2015

	Profitability			Trades		
	P-T	Broker A	Broker B	P-T	Broker A	Broker B
Total	\$ 4,868,587	\$ 3,956,693	\$ 4,082,885	4,248,460	3,942,358	2,516,189
Trading	\$ 3,831,375	\$ 3,077,551	\$ 3,631,226	4,248,460	3,942,358	2,516,189
Dly avg.	\$ 31,928	\$ 25,646	\$ 30,260	35,404	32,853	20,968
Dly Stddev.	\$ 19,057	\$ 18,073	\$ 15,349	9,098	8,772	6,555
Dly Median	\$ 29,698	\$ 23,776	\$ 27,369	34,164	31,980	19,676
Dly min.	-\$ 3,513	-\$ 3,422	-\$ 480	8,492	7,487	4,126
1-min avg.	\$ 82.50	\$ 66.27	\$ 78.19	91	85	54
1-min stddev.	\$ 485.48	\$ 561.58	\$ 460.96	56	55	38
1-min median	\$ 40.65	\$ 28.65	\$ 38.99	77	70	44
1-min min.	-\$ 7,846.67	-\$ 14,210.84	-\$ 13,870.98	4	2	0

This table shows aggregated statistics on the two prioritization scenarios: price time (P-T), the international standard, and broker preference applied to Brokers A and B. Total profitability is the cumulative trading profit including fees and rebates over the full sample. Trading excludes fees and rebates. Total trade is the number of trades executed by the ELP. The abbreviation significance is: dly: daily, avg: average, stddev: standard deviation, min: minimum, and 1-min: one-minute time interval.

Table 3. 5 Monthly performance : price-time and Brokers A and B

		(1)	(2)	(3)	(4)	(5)	(6)
	Month	Trading	Rebate	Net	(2)/(3)	# trades	# LO/# tr
P.T.	March	\$ 722,265	\$ 242,182	\$ 964,447	25.1%	777,643	5.56
P.T.	April	\$ 521,544	\$ 222,620	\$ 744,163	29.9%	693,597	5.81
P.T.	May	\$ 315,995	\$ 134,555	\$ 450,550	29.9%	434,733	5.68
P.T.	June	\$ 697,936	\$ 126,600	\$ 824,536	15.4%	706,331	5.62
P.T.	July	\$ 672,249	\$ 136,907	\$ 809,156	16.9%	734,707	5.44
P.T.	August	\$ 901,387	\$ 174,348	\$ 1,075,735	16.2%	901,449	5.33
		\$ 3,831,375	\$ 1,037,212	\$ 4,868,588	21.3%	4,248,460	5.57
A	March	\$ 530,897	\$ 217,526	\$ 748,423	29.1%	741,070	5.14
A	April	\$ 358,125	\$ 190,386	\$ 548,512	34.7%	635,199	5.20
A	May	\$ 214,780	\$ 116,620	\$ 331,400	35.2%	405,375	4.96
A	June	\$ 570,076	\$ 105,555	\$ 675,631	15.6%	644,849	4.89
A	July	\$ 521,699	\$ 110,389	\$ 632,088	17.5%	667,103	4.62
A	August	\$ 881,973	\$ 138,667	\$ 1,020,640	13.6%	848,762	4.54
		\$ 3,077,551	\$ 879,142	\$ 3,956,693	22.2%	3,942,358	4.89
B	March	\$ 663,623	\$ 102,112	\$ 765,735	13.3%	439,007	3.71
B	April	\$ 474,864	\$ 89,431	\$ 564,295	15.8%	372,434	3.75
B	May	\$ 307,986	\$ 54,014	\$ 362,000	14.9%	240,863	3.58
B	June	\$ 645,858	\$ 56,356	\$ 702,214	8.0%	429,835	3.76
B	July	\$ 663,220	\$ 62,612	\$ 725,832	8.6%	449,343	3.68
B	August	\$ 875,674	\$ 87,135	\$ 962,809	9.1%	584,707	3.90
		\$ 3,631,226	\$ 451,659	\$ 4,082,885	11.1%	2,516,189	3.73

This table provides monthly statistics from price time (P.T.) and Brokers A and B activities. Column (1) shows the gross profit from trading; column (2) is the net result of maker rebates, taker fees and IIROC fees; column (3) is column (1) minus column (2); column (4) is the number of ELP trades; column (5) is the ratio of the number of ELP limit orders executed against incoming marketable orders over the number of ELP executed marketable orders.

Table 3. 6 Stable distribution parameters : Brokers A and B profits & losses March – August 2015

	A	B
Alpha	1.2823	1.2615
Beta	0.2779	0.2185
Gamma	148.4979	178.3119
Delta	117.7321	108.4078

This table supplies the parameters value from the P&L distributions for Brokers A and B under the broker preference scenario. The stable family distributions are described by four parameters: $\alpha \in (0,2]$: a stability parameter; $\beta \in [-1,1]$: a skewness parameter; $\gamma \in (0, \infty)$: a scale parameter; $\delta \in (-\infty, +\infty)$: a location parameter.

Table 3. 7 Market maker index: Brokers A and B

	# buy limit order	# sell limit order	ratio
A	13,395,789,249	18,902,619,467	0.1666
B	12,145,242,039	18,129,883,345	0.1986

This table provides statistics concerning the Brokers A and B buy and sell limit orders. # is the abbreviation of number. The ratios are the absolute value of market maker index obtained from Equation ().

Table 3. 8 Quotes and trades : Brokers A and B

		(1)	(2)	(3)	(4)	(5)	(6)
		Quotes			Trades		
		Actual	ELP	%	Actual	ELP	%
March	A	107,670,190	18,224,935	16.93%	6,244,119	741,070	11.87%
April	A	99,615,388	15,342,662	15.40%	5,569,695	635,199	11.40%
May	A	61,020,594	10,165,434	16.66%	3,682,179	405,375	11.01%
June	A	82,540,777	15,109,659	18.31%	5,743,586	644,849	11.23%
July	A	84,090,062	14,707,922	17.49%	6,118,321	667,103	10.90%
August	A	106,355,892	19,745,589	18.57%	7,687,282	848,762	11.04%
		541,292,903	93,296,201	17.24%	35,045,182	3,942,358	11.25%
March	B	107,670,190	18,197,298	16.90%	6,244,119	439,007	7.03%
April	B	99,615,388	15,427,856	15.49%	5,569,695	372,434	6.69%
May	B	61,020,594	10,090,060	16.54%	3,682,179	240,863	6.54%
June	B	82,540,777	14,853,049	17.99%	5,743,586	429,835	7.48%
July	B	84,090,062	14,505,654	17.25%	6,118,321	449,343	7.34%
August	B	106,355,892	19,614,464	18.44%	7,687,282	584,707	7.61%
		541,292,903	92,688,381	17.12%	35,045,182	2,516,189	7.18%

This table compares ELP to actual (observed) trades and quotes. Column (3) and column (6) are the percentages of ELP activities during the simulation over the actual ones.

Table 3. 9 Regressions on factors determining ELP's variations of position

Broker A	(1)	(2)	(3)	(4)	(5)	(6)
$\text{diffMV}_{i,d,t}$	0.09*** (3.13)	0.09*** (3.29)	0.09*** (3.29)	0.07*** (2.67)	0.07*** (2.67)	0.07*** (2.68)
$\text{pct}\sigma_{i,d,t}$		404,347*** (6.10)	353,761*** (5.98)	353,761*** (6.10)	353,761*** (6.10)	353,765*** (6.10)
$\text{diffSpread}_{i,d,t}$			34,561** (2.32)	33,302** (2.25)	33,302** (2.25)	33,302** (2.25)
$\text{diff}\#\text{Quotes}_{i,d,t}$				61.43*** (6.24)	61.43*** (6.24)	61.43*** (6.24)
$\text{diffImbal}_{i,d,t}$					0.00 (0.85)	0.00 (0.85)
$\text{pctMom}_{m,d,t}$						-0.05 (0.11)
R ²	3.95%	4.60%	4.69%	5.44%	5.44%	5.46%
N. obs.	2,725,800	2,725,800	2,725,800	2,725,800	2,725,800	2,725,800
Broker B	(1)	(2)	(3)	(4)	(5)	(6)
$\text{diffMV}_{i,d,t}$	0.05*** (4.31)	0.04*** (4.29)	0.04*** (4.30)	0.04*** (3.43)	0.04*** (3.43)	0.04*** (3.43)
$\text{pct}\sigma_{i,d,t}$		153,545*** (4.07)	151,254*** (4.01)	125,911*** (4.14)	125,911*** (4.14)	125,930*** (4.14)
$\text{diffSpread}_{i,d,t}$			15,647*** (3.14)	14,946*** (3.05)	14,946*** (3.05)	14,946*** (3.05)
$\text{diff}\#\text{Quotes}_{i,d,t}$				34.20*** (6.77)	34.20*** (6.77)	34.20*** (6.77)
$\text{diffImbal}_{i,d,t}$					0.00 (0.50)	0.00 (0.50)
$\text{pctMom}_{m,d,t}$						-0.19 (1.20)
R ²	3.82%	4.18%	4.26%	5.16%	5.16%	5.16%
N. obs.	2,725,800	2,725,800	2,725,800	2,725,800	2,725,800	2,725,800

The table represents the results from estimating Equation () for the full sample of the S&P/TSX 60 Composite securities. This panel regression is based on per stock and one-minute time intervals. $\text{diffMV}_{i,d,t}$ is the first difference in total market value traded for stock i on day d during time interval t , a measure of trading intensity. $\text{pct}\sigma_{i,d,t}$ is the stock i standard deviation relative performance on day d during the time interval t , a relative measure of idiosyncratic risk direction. $\text{diffSprd}_{i,d,t}$ is the first difference in bid-ask spread for stock i on day d at the end of the time interval t , a measure of ELP remuneration. $\text{diffImbal}_{i,d,t}$ is the first difference in LOB level 1 imbalance for stock i on day d at the end of the time interval t , a liquidity disequilibrium measure. $\text{diff}\#\text{Quotes}_{i,d,t}$ is the first difference in the number of quotes for stock i on day d during time interval t , a measure of liquidity competition. $\text{pctMom}_{m,d,t}$ is the relative performance of the maximum between the proportion of up and down volume for the market on day d at the end of interval t , a measure of relative momentum. The underlying regressions include stock standard errors double-clustered by security and time interval. *, **, *** indicate significance at 10%, 5%, and 1% levels; the table shows the t-statistic. Variations of position for Brokers A and B are analyzed.

Table 3. 10 Broker B Level 1 bid (upper part of) and offer (lower part of) cross correlation matrix during the extreme events

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	1.00	0.36	0.24	0.22	0.11	0.10	0.17	0.00	0.20	0.29	0.15	0.08	0.20	0.27	0.13	-0.01	0.12	0.13	0.24	0.02	0.21	0.13	0.01	0.11	0.14	0.27	0.19	0.39	0.18	0.22
2	0.44	1.00	0.10	0.22	0.02	0.11	0.13	0.00	0.27	0.28	0.26	0.12	0.18	0.22	0.12	0.15	0.21	0.21	0.23	0.05	0.10	0.14	0.05	0.21	0.19	0.34	0.17	0.42	0.18	0.16
3	0.34	0.28	1.00	0.04	0.18	0.20	0.12	0.00	0.16	0.29	0.19	0.09	0.26	0.14	0.11	0.06	0.05	0.16	0.19	0.06	0.05	0.06	0.06	0.16	0.12	0.09	0.18	0.16	0.15	0.15
4	0.20	0.10	0.10	1.00	0.13	0.26	0.18	0.00	0.22	0.30	0.28	0.26	0.23	0.30	0.14	0.08	0.28	0.21	0.26	0.00	0.22	0.10	0.01	0.30	0.23	0.19	0.18	0.30	0.17	0.15
5	0.11	0.04	0.12	0.12	1.00	0.21	0.09	0.00	0.18	0.13	0.17	0.14	0.21	0.15	0.23	0.05	0.07	0.16	0.05	0.02	0.05	0.02	-0.05	0.19	0.11	0.03	0.17	0.06	0.22	0.16
6	0.16	0.18	0.28	0.15	0.17	1.00	0.13	0.00	0.23	0.36	0.33	0.18	0.31	0.22	0.19	0.13	0.23	0.21	0.22	-0.06	0.11	0.15	0.04	0.25	0.09	0.13	0.14	0.16	0.20	0.18
7	0.32	0.29	0.21	0.09	0.13	0.19	1.00	0.00	0.11	0.23	0.19	0.12	0.12	0.14	0.16	0.05	0.13	0.13	0.23	0.02	0.09	0.20	0.14	0.16	0.13	0.08	0.21	0.19	0.11	0.17
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.28	0.32	0.21	0.18	0.26	0.28	0.22	0.00	1.00	0.34	0.40	0.23	0.33	0.32	0.33	0.19	0.25	0.19	0.25	0.05	0.23	0.01	0.05	0.33	0.21	0.27	0.19	0.35	0.18	0.17
10	0.35	0.39	0.25	0.21	0.22	0.29	0.34	0.00	0.42	1.00	0.44	0.22	0.39	0.28	0.26	0.16	0.33	0.27	0.31	0.06	0.16	0.11	0.01	0.30	0.27	0.28	0.22	0.31	0.26	0.22
11	0.24	0.31	0.16	0.13	0.21	0.18	0.26	0.00	0.39	0.39	1.00	0.29	0.43	0.28	0.29	0.19	0.22	0.23	0.36	0.03	0.07	0.05	0.12	0.30	0.25	0.23	0.21	0.26	0.14	0.09
12	0.37	0.28	0.25	0.35	0.17	0.26	0.26	0.00	0.40	0.45	0.24	1.00	0.24	0.27	0.26	0.10	0.15	0.19	0.22	0.00	0.08	0.06	0.01	0.21	0.20	0.10	0.18	0.17	0.16	0.16
13	0.26	0.35	0.24	0.08	0.17	0.28	0.25	0.00	0.37	0.46	0.40	0.24	1.00	0.23	0.18	0.17	0.14	0.28	0.29	0.11	0.10	0.05	0.05	0.21	0.28	0.31	0.29	0.32	0.24	0.18
14	0.29	0.29	0.18	0.38	0.21	0.19	0.27	0.00	0.32	0.40	0.29	0.42	0.17	1.00	0.23	0.10	0.29	0.18	0.31	0.09	0.26	0.09	0.00	0.27	0.27	0.21	0.25	0.31	0.23	0.19
15	0.21	0.25	0.21	0.08	0.12	0.23	0.22	0.00	0.37	0.40	0.31	0.30	0.38	0.15	1.00	0.15	0.25	0.10	0.23	0.04	0.17	0.01	-0.01	0.23	0.18	0.13	0.15	0.16	0.17	0.18
16	-0.05	0.07	0.04	0.08	0.15	0.10	0.04	0.00	0.10	0.17	0.12	0.07	0.10	0.16	0.14	1.00	0.11	0.14	0.07	0.05	0.02	0.02	0.05	0.12	0.06	0.11	0.04	0.17	0.05	0.08
17	0.22	0.13	0.15	0.31	0.17	0.21	0.18	0.00	0.23	0.20	0.23	0.28	0.13	0.37	0.16	0.03	1.00	0.21	0.23	0.06	0.23	0.09	-0.05	0.24	0.22	0.21	0.08	0.23	0.19	0.18
18	0.17	0.15	0.11	0.12	0.18	0.07	0.10	0.00	0.12	0.21	0.21	0.04	0.19	0.08	0.09	0.17	0.06	1.00	0.21	0.23	0.15	0.11	0.01	0.22	0.16	0.15	0.18	0.20	0.16	0.10
19	0.27	0.22	0.20	0.32	0.15	0.17	0.27	0.00	0.34	0.36	0.28	0.40	0.29	0.36	0.26	0.03	0.25	0.08	1.00	0.06	0.30	0.10	0.08	0.23	0.23	0.27	0.22	0.32	0.14	0.16
20	0.07	0.09	0.10	0.03	0.05	0.08	0.02	0.00	0.08	0.06	0.11	0.10	0.04	0.09	0.00	0.06	0.01	0.11	0.18	1.00	0.06	0.03	0.04	-0.04	0.02	0.11	-0.03	0.05	0.09	0.04
21	0.30	0.12	0.11	0.30	0.15	0.14	0.18	0.00	0.16	0.21	0.12	0.35	0.04	0.32	0.10	0.00	0.30	0.06	0.32	0.04	1.00	0.06	0.04	0.13	0.15	0.21	0.15	0.20	0.11	0.19
22	0.35	0.19	0.13	0.08	0.12	0.06	0.14	0.00	0.07	0.10	0.10	0.19	0.16	0.15	0.06	-0.11	0.21	0.09	0.18	0.01	0.16	1.00	0.06	0.11	0.14	0.06	0.14	0.19	0.11	0.08
23	0.05	0.12	0.11	0.14	0.10	0.02	0.12	0.00	0.10	0.08	0.17	0.08	0.10	0.18	0.03	0.03	0.08	0.11	0.15	0.08	0.05	0.01	1.00	0.01	0.00	0.01	0.07	-0.06	0.00	-0.03
24	0.28	0.24	0.27	0.22	0.13	0.26	0.17	0.00	0.32	0.30	0.27	0.31	0.26	0.26	0.31	0.09	0.19	0.10	0.26	0.05	0.17	0.09	0.09	1.00	0.13	0.15	0.26	0.24	0.15	0.19
25	0.16	0.21	0.05	0.23	0.06	0.08	0.15	0.00	0.20	0.29	0.11	0.25	0.12	0.34	0.21	0.15	0.26	0.12	0.24	0.04	0.25	0.07	0.06	0.22	1.00	0.21	0.18	0.26	0.12	0.14
26	0.40	0.47	0.21	0.13	0.04	0.20	0.23	0.00	0.24	0.29	0.23	0.30	0.29	0.18	0.28	0.11	0.15	0.10	0.25	0.12	0.15	0.20	0.02	0.25	0.21	1.00	0.16	0.41	0.14	0.24
27	0.10	0.10	0.09	0.33	0.26	0.19	0.06	0.00	0.34	0.21	0.29	0.26	0.15	0.30	0.20	0.13	0.24	0.16	0.24	0.03	0.17	0.11	0.09	0.25	0.13	0.10	1.00	0.21	0.16	0.16
28	0.55	0.48	0.37	0.24	0.14	0.23	0.26	0.00	0.30	0.45	0.27	0.39	0.32	0.33	0.28	0.02	0.24	0.17	0.34	0.11	0.29	0.28	0.08	0.30	0.34	0.41	0.11	1.00	0.21	0.32
29	0.32	0.30	0.29	0.27	0.22	0.27	0.25	0.00	0.40	0.40	0.33	0.36	0.32	0.35	0.23	0.10	0.32	0.19	0.26	0.01	0.24	0.15	0.13	0.27	0.29	0.29	0.35	0.33	1.00	0.21
30	0.31	0.29	0.29	0.14	0.16	0.20	0.30	0.00	0.34	0.30	0.26	0.32	0.33	0.26	0.32	0.00	0.27	0.06	0.27	0.01	0.15	0.17	0.10	0.32	0.17	0.27	0.12	0.32	0.33	1.00
31	0.26	0.18	0.21	0.32	0.18	0.13	0.11	0.00	0.22	0.27	0.26	0.34	0.14	0.38	0.16	0.19	0.32	0.23	0.32	0.08	0.31	0.12	0.17	0.27	0.24	0.20	0.27	0.34	0.36	0.14
32	0.21	0.27	0.19	0.21	0.18	0.19	0.20	0.00	0.28	0.24	0.25	0.37	0.29	0.27	0.27	0.00	0.21	0.09	0.33	0.05	0.22	0.11	0.08	0.17	0.18	0.24	0.24	0.31	0.33	0.29
33	0.17	0.22	0.20	0.31	0.26	0.15	0.15	0.00	0.31	0.26	0.31	0.31	0.24	0.29	0.26	0.11	0.28	0.09	0.28	0.05	0.24	0.14	0.07	0.30	0.30	0.24	0.29	0.29	0.32	0.31
34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
35	0.14	0.12	0.09	0.23	0.20	0.16	0.13	0.00	0.25	0.31	0.23	0.22	0.17	0.25	0.18	0.13	0.14	0.19	0.26	0.09	0.16	0.03	0.08	0.29	0.24	0.16	0.28	0.23	0.24	0.19
36	0.31	0.30	0.26	0.20	0.20	0.25	0.29	0.00	0.34	0.47	0.36	0.34	0.31	0.27	0.41	0.06	0.22	0.15	0.31	0.00	0.15	0.16	0.06	0.29	0.18	0.27	0.13	0.34	0.30	0.30
37	0.33	0.32	0.31	0.25	0.17	0.29	0.19	0.00	0.36	0.40	0.28	0.45	0.32	0.33	0.37	0.13	0.29	0.14	0.30	0.03	0.22	0.11	0.04	0.37	0.28	0.32	0.24	0.42	0.39	0.32
38	0.12	0.20	0.11	0.12	0.17	0.16	0.19	0.00	0.26	0.30	0.26	0.13	0.25	0.20	0.25	0.18	0.12	0.20	0.16	0.10	0.04	0.02	0.05	0.16	0.16	0.18	0.23	0.13	0.22	0.17
39	0.29	0.36	0.27	0.25	0.17	0.19	0.24	0.00	0.38	0.42	0.35	0.35	0.34	0.36	0.37	0.08	0.18	0.14	0.34	0.08	0.26	0.16	0.09	0.31	0.20	0.31	0.28	0.36	0.37	0.38
40	0.42	0.41	0.23	0.09	0.12	0.19	0.42	0.00	0.34	0.48	0.29	0.31	0.34	0.28	0.39	0.04	0.17	0.08	0.27	0.03	0.13	0.1								

Table 3. 11 Broker B Level 1 bid (upper part of) and offer (lower part of) cross correlation matrix during the extreme events

	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	Avg
1	0.17	0.19	0.10	0.00	0.09	0.15	0.16	0.18	0.22	0.24	0.14	0.12	0.18	0.19	0.17	0.15	0.28	0.22	0.32	0.15	0.21	0.09	0.12	0.24	0.18	0.26	0.01	0.06	0.20	0.17
2	0.18	0.20	0.17	0.00	0.10	0.20	0.22	0.10	0.22	0.32	0.16	0.22	0.31	0.13	0.22	0.00	0.27	0.23	0.33	0.14	0.19	0.24	0.12	0.26	0.27	0.25	-0.01	0.12	0.08	0.18
3	0.08	0.16	0.04	0.00	0.13	0.13	0.22	0.21	0.17	0.16	0.16	0.16	0.20	0.15	0.21	0.17	0.16	0.13	0.13	0.08	0.21	0.14	0.07	0.21	0.18	0.30	0.08	0.09	0.08	0.14
4	0.24	0.25	0.19	0.00	0.16	0.29	0.24	0.10	0.34	0.28	0.19	0.27	0.28	0.22	0.25	0.13	0.21	0.23	0.38	0.19	0.35	0.19	0.19	0.27	0.21	0.28	0.06	0.09	0.18	0.20
5	0.11	0.13	0.26	0.00	0.23	0.15	0.12	0.20	0.12	0.07	0.22	0.16	0.15	0.06	0.22	0.10	0.19	0.15	0.10	0.11	0.08	0.25	0.01	0.11	0.15	0.20	0.07	0.07	0.06	0.12
6	0.07	0.28	0.15	0.00	0.15	0.20	0.36	0.18	0.29	0.30	0.18	0.27	0.29	0.21	0.31	0.15	0.24	0.28	0.29	0.14	0.25	0.30	0.25	0.29	0.21	0.32	0.07	0.03	0.08	0.19
7	0.20	0.17	0.15	0.00	0.13	0.18	0.12	0.14	0.14	0.26	-0.05	0.07	0.17	0.21	0.11	0.05	0.14	0.17	0.27	0.13	0.23	0.11	0.14	0.19	0.23	0.21	-0.02	0.11	0.21	0.15
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.25	0.18	0.25	0.00	0.22	0.39	0.22	0.12	0.34	0.29	0.30	0.30	0.19	0.17	0.37	0.19	0.28	0.37	0.33	0.08	0.37	0.26	0.13	0.42	0.35	0.29	0.04	0.13	0.06	0.23
10	0.24	0.32	0.27	0.00	0.20	0.33	0.31	0.14	0.34	0.39	0.20	0.29	0.32	0.28	0.34	0.30	0.35	0.40	0.39	0.29	0.42	0.25	0.24	0.46	0.33	0.37	0.01	0.09	0.14	0.26
11	0.16	0.25	0.24	0.00	0.27	0.41	0.34	0.21	0.41	0.35	0.19	0.38	0.21	0.17	0.46	0.11	0.26	0.45	0.28	0.22	0.32	0.28	0.18	0.45	0.26	0.36	0.01	0.16	0.10	0.24
12	0.19	0.15	0.19	0.00	0.20	0.23	0.16	0.10	0.23	0.22	0.17	0.25	0.20	0.29	0.23	0.17	0.17	0.28	0.24	0.10	0.28	0.16	0.19	0.26	0.20	0.22	0.06	0.00	0.06	0.19
13	0.18	0.18	0.17	0.00	0.27	0.31	0.26	0.15	0.28	0.26	0.28	0.23	0.30	0.24	0.44	0.13	0.33	0.41	0.27	0.18	0.33	0.24	0.08	0.45	0.26	0.37	0.04	0.10	0.15	0.23
14	0.28	0.24	0.17	0.00	0.11	0.24	0.20	0.11	0.24	0.23	0.28	0.32	0.30	0.26	0.23	0.12	0.29	0.29	0.35	0.13	0.38	0.09	0.22	0.27	0.28	0.32	0.05	0.04	0.19	0.22
15	0.15	0.20	0.22	0.00	0.11	0.27	0.22	0.13	0.30	0.33	0.21	0.27	0.20	0.17	0.27	0.15	0.13	0.28	0.19	0.13	0.28	0.17	0.13	0.30	0.26	0.27	0.10	0.07	0.10	0.19
16	0.06	0.14	0.01	0.00	-0.02	0.09	0.18	0.10	0.18	0.09	0.16	0.19	0.14	0.02	0.29	0.07	0.01	0.17	0.05	0.05	0.07	0.12	-0.02	0.17	0.10	0.12	0.06	0.08	-0.07	0.09
17	0.19	0.18	0.26	0.00	0.17	0.22	0.18	0.08	0.21	0.26	0.14	0.28	0.21	0.14	0.17	0.08	0.19	0.17	0.28	0.15	0.35	0.16	0.18	0.23	0.23	0.23	0.00	0.07	0.12	0.18
18	0.17	0.16	0.22	0.00	0.24	0.26	0.10	0.13	0.15	0.21	0.21	0.13	0.18	0.19	0.24	0.15	0.24	0.19	0.23	0.11	0.25	0.21	0.10	0.20	0.31	0.23	-0.05	0.14	0.12	0.15
19	0.34	0.28	0.17	0.00	0.25	0.25	0.15	0.16	0.27	0.34	0.11	0.26	0.27	0.20	0.24	0.10	0.25	0.37	0.38	0.18	0.34	0.19	0.20	0.29	0.25	0.29	0.04	0.13	0.17	0.22
20	0.09	-0.05	0.06	0.00	0.10	0.02	0.01	-0.06	-0.02	-0.04	0.04	0.00	0.03	0.03	0.04	0.05	0.15	0.05	0.10	-0.01	0.09	0.07	-0.03	0.03	0.10	0.06	-0.01	-0.01	0.03	0.04
21	0.25	0.10	0.14	0.00	0.09	0.15	0.16	0.11	0.17	0.13	0.11	0.10	0.10	0.16	0.11	0.19	0.12	0.16	0.23	0.12	0.19	0.04	0.10	0.17	0.17	0.09	0.03	0.05	0.09	0.14
22	0.10	0.15	0.04	0.00	0.15	0.12	0.10	0.09	0.08	0.13	0.07	0.11	0.07	0.26	0.07	0.10	0.13	0.10	0.16	0.07	0.10	0.05	0.11	0.10	0.12	0.18	0.06	0.00	0.19	0.11
23	0.03	0.10	-0.02	0.00	0.03	0.03	-0.01	0.00	0.03	0.10	-0.03	0.04	-0.03	-0.01	0.07	0.07	-0.04	0.03	-0.08	0.06	-0.05	-0.03	-0.02	0.10	0.04	-0.02	-0.02	0.07	-0.05	0.04
24	0.14	0.23	0.21	0.00	0.20	0.31	0.28	0.19	0.23	0.28	0.29	0.33	0.21	0.22	0.31	0.16	0.25	0.34	0.23	0.16	0.31	0.22	0.16	0.30	0.24	0.25	0.06	0.13	0.06	0.21
25	0.20	0.11	0.11	0.00	0.16	0.21	0.11	0.16	0.17	0.22	0.18	0.23	0.27	0.16	0.18	0.17	0.19	0.30	0.23	0.12	0.27	0.05	0.17	0.19	0.19	0.24	0.06	0.00	0.20	0.27
26	0.15	0.13	0.16	0.00	0.13	0.14	0.20	0.08	0.31	0.20	0.14	0.20	0.24	0.07	0.18	0.07	0.20	0.21	0.28	0.15	0.28	0.11	0.06	0.28	0.22	0.25	0.05	0.02	0.14	0.18
27	0.29	0.14	0.17	0.00	0.19	0.19	0.16	0.24	0.16	0.14	0.20	0.16	0.20	0.19	0.21	0.11	0.25	0.28	0.17	0.21	0.23	0.09	0.16	0.22	0.17	0.16	-0.01	0.03	0.09	0.17
28	0.24	0.25	0.23	0.00	0.11	0.25	0.22	0.14	0.31	0.32	0.22	0.21	0.31	0.21	0.24	0.10	0.32	0.31	0.40	0.18	0.31	0.13	0.14	0.34	0.30	0.28	0.06	0.04	0.24	0.25
29	0.19	0.24	0.19	0.00	0.12	0.19	0.24	0.09	0.14	0.19	0.22	0.14	0.23	0.19	0.22	0.15	0.22	0.19	0.20	0.13	0.29	0.10	0.16	0.28	0.21	0.26	0.10	0.06	0.06	0.21
30	0.26	0.23	0.15	0.00	0.08	0.14	0.21	0.17	0.13	0.14	0.09	0.07	0.25	0.26	0.16	0.07	0.24	0.17	0.20	0.17	0.23	0.08	0.13	0.22	0.20	0.14	0.13	-0.02	0.07	0.19
31	1.00	0.16	0.22	0.00	0.12	0.20	0.12	0.11	0.11	0.17	0.13	0.12	0.24	0.16	0.11	0.20	0.23	0.20	0.23	0.15	0.26	0.10	0.12	0.21	0.22	0.18	0.05	0.08	0.09	0.19
32	0.26	1.00	0.22	0.00	0.12	0.20	0.31	0.23	0.26	0.36	0.16	0.32	0.27	0.26	0.28	0.12	0.19	0.19	0.23	0.23	0.30	0.17	0.25	0.23	0.21	0.31	0.16	0.08	0.08	0.21
33	0.28	0.31	1.00	0.00	0.17	0.22	0.16	0.15	0.30	0.19	0.28	0.21	0.21	0.16	0.21	0.10	0.16	0.20	0.17	0.22	0.26	0.20	0.15	0.20	0.22	0.22	0.00	0.17	0.10	0.21
34	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
35	0.25	0.15	0.28	0.00	1.00	0.16	0.14	0.21	0.16	0.07	0.18	0.15	0.13	0.16	0.25	0.13	0.22	0.26	0.18	0.09	0.17	0.19	0.03	0.21	0.19	0.19	-0.02	0.08	0.03	0.17
36	0.15	0.24	0.29	0.00	0.25	1.00	0.19	0.09	0.29	0.31	0.24	0.30	0.18	0.23	0.31	0.15	0.25	0.36	0.30	0.16	0.40	0.17	0.24	0.36	0.37	0.23	0.04	0.16	0.15	0.23
37	0.30	0.25	0.33	0.00	0.18	0.38	1.00	0.19	0.36	0.30	0.20	0.28	0.25	0.16	0.29	0.19	0.15	0.28	0.23	0.18	0.26	0.17	0.16	0.30	0.18	0.27	0.16	0.05	0.07	0.24
38	0.14	0.09	0.24	0.00	0.29	0.26	0.20	1.00	0.13	0.06	0.20	0.16	0.16	0.20	0.16	0.17	0.13	0.22	0.09	0.26	0.17	0.10	0.15	0.09	0.15	0.13	0.08	0.02	0.16	0.16
39	0.25	0.31	0.39	0.00	0.28	0.34	0.31	0.25	1.00	0.26	0.29	0.27	0.29	0.19	0.35	0.11	0.25	0.30	0.23	0.21	0.24	0.16	0.13	0.32	0.27	0.28	0.05	0.12	0.18	0.25
40	0.13	0.22	0.17	0.00	0.16	0.44	0.40	0.15	0.28	1.00	0.10	0.33	0.26	0.13	0.28	0.20	0.19	0.32	0.36	0.20	0.37	0.17	0.30	0.36	0.38	0.33	0.05	0.16	0.14	0.23
41	0.19	0.22	0.21	0.00	0.10	0.20	0.28	0.18	0.25	0.19	1.00	0.21	0.19	0.13	0.30	0.16	0.22	0.30	0.08	0.21	0.25	0.09	0.09	0.18	0.20	0.22	0.08	0.02	0.03	0.17
42	0.18	0.29	0.31	0.00	0.22	0.30	0.26	0.21	0.25	0.27	0.25	1.00	0.26	0.17	0.32	0.11	0.12	0.33	0.25	0.18	0.27	0.25	0.21	0.30	0.23	0.30	0.09	0.13	0.09	0.22
43	0.31	0.27	0.24	0.00	0.27	0.29	0.35	0.13	0.33	0.28	0.21	0.23	1.00	0.19	0.25	0.09	0.26	0.23	0.37	0.15	0.30	0.17	0.19	0.29						

Figure 3. 1 VIX S&P/TSX 60, XIU.TO : March - August 2015

This graph supplies the behavior of the implicit volatility from the VIX S&P/TSX 60 (left scale) and the prices from XIU.TO, an ETF tracking the S&P/TSX60 (right scale).

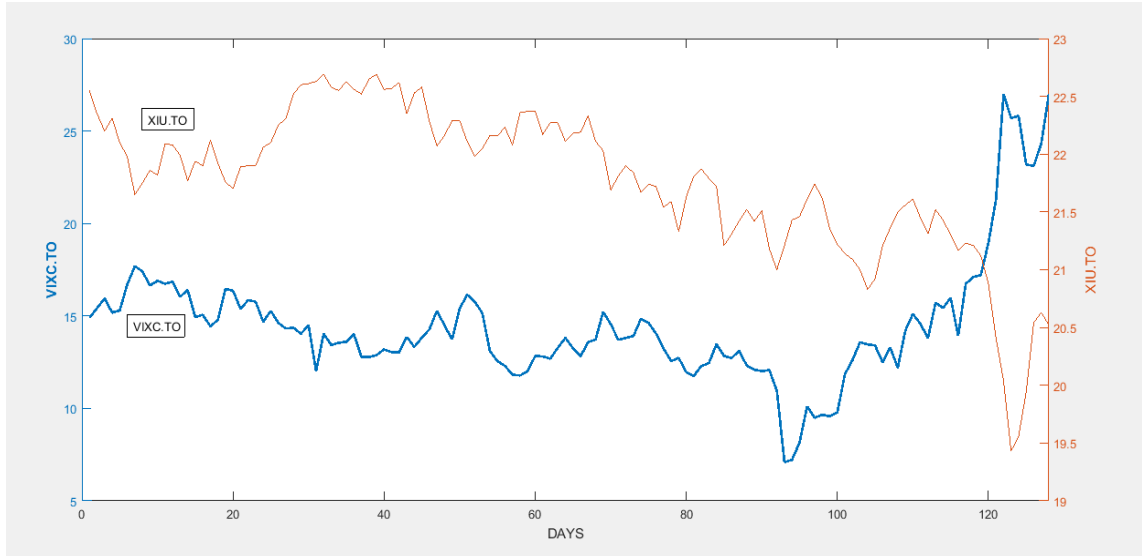
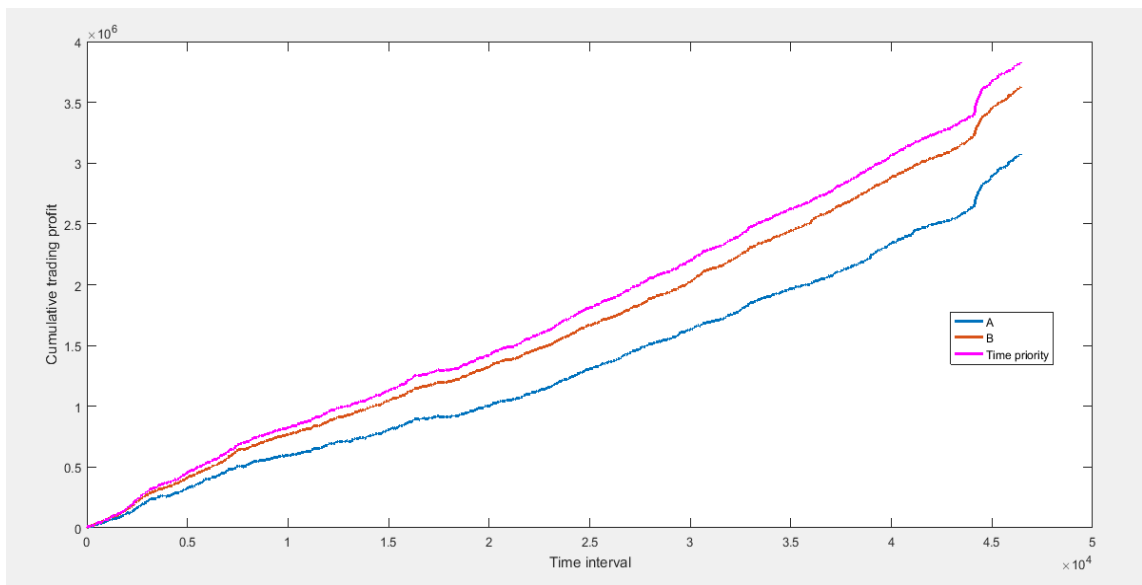
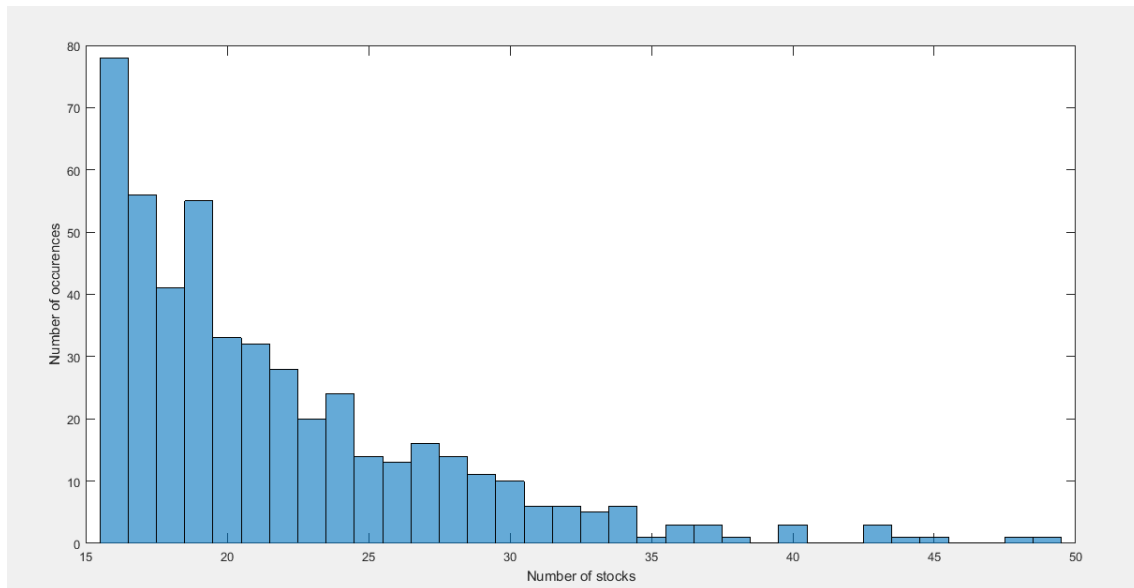


Figure 3. 2 Cumulative profit excluding rebates: broker preference with Brokers A and B vs. time-price priority



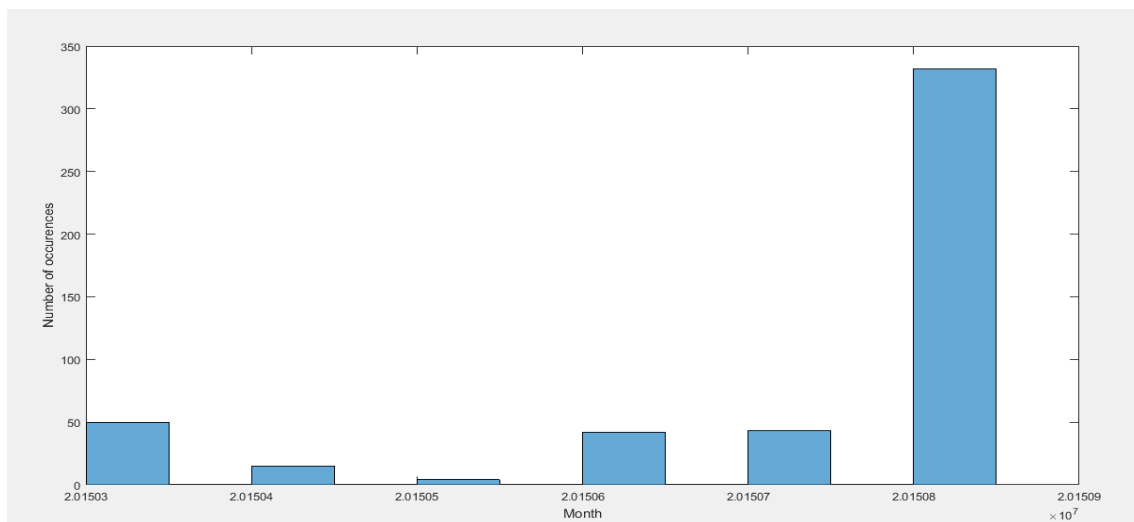
This graph presents the cumulative trading profit excluding fees and rebates. It compares the broker preference with results obtained by Brokers A (the lower line) and B (the middle line) to the international standard in prioritization based on price and time. Time intervals are of one minute. Profits are aggregated over the traded S&P/TSX60 components.

Figure 3. 3 Critical market states: repartition by number of stocks involved : March – August 2015



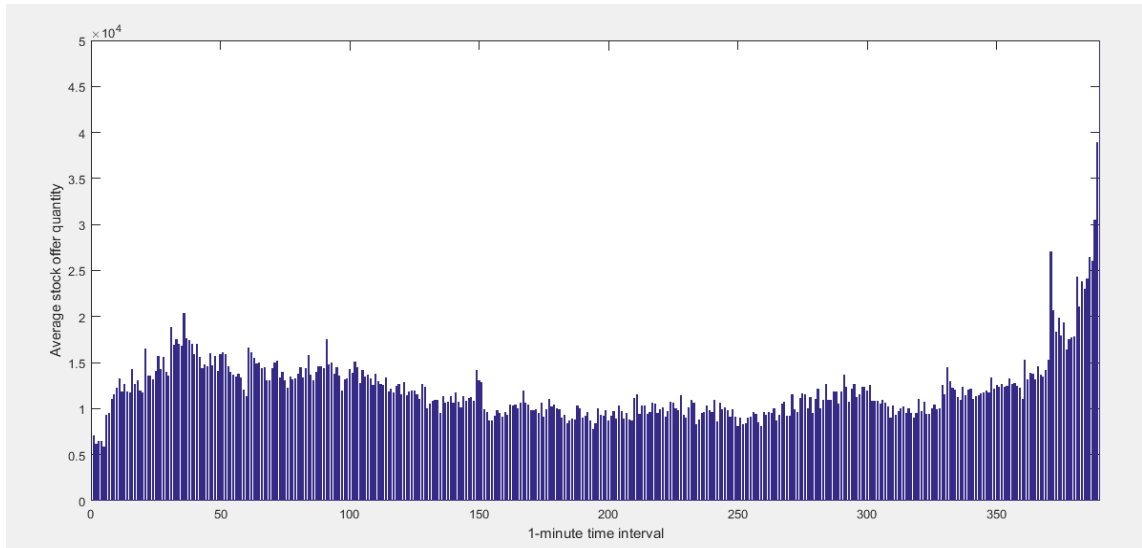
This histogram reports the number of occurrences of critical market states grouped by the number of stocks that the absolute logarithmic return and the number of trades exceed simultaneously their quantile values of 90%.

Figure 3. 4 Critical market states: repartition by monthly occurrences : March – August 2015



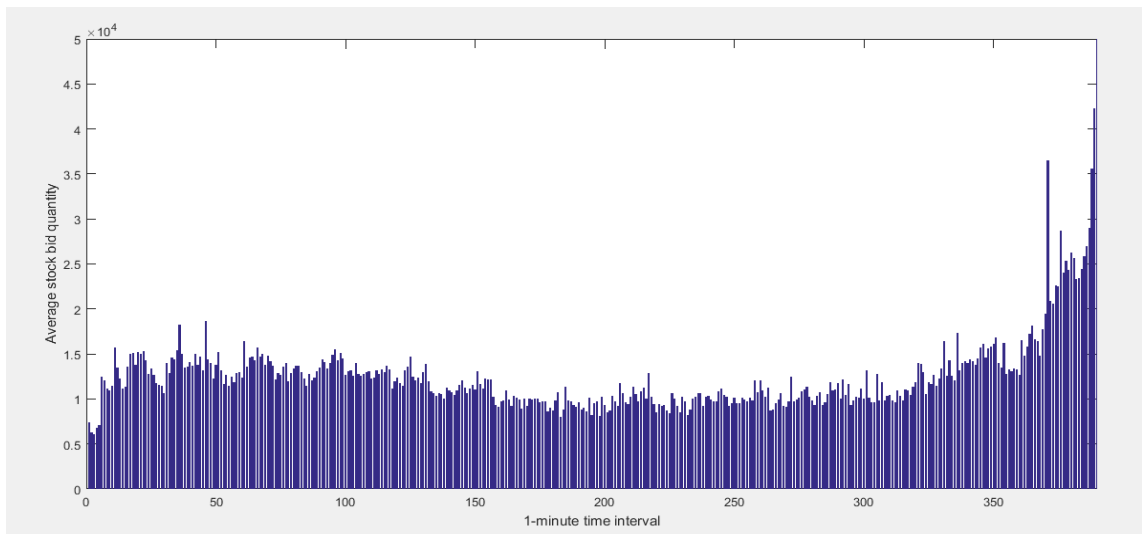
This histogram depicts the number of monthly occurrences of critical market states. The market is in a critical state when a minimum of 15 stocks simultaneously exceeds the quantile values of 90% for their absolute logarithmic return and their number of trades.

Figure 3. 5 Average level 1 offer quantity



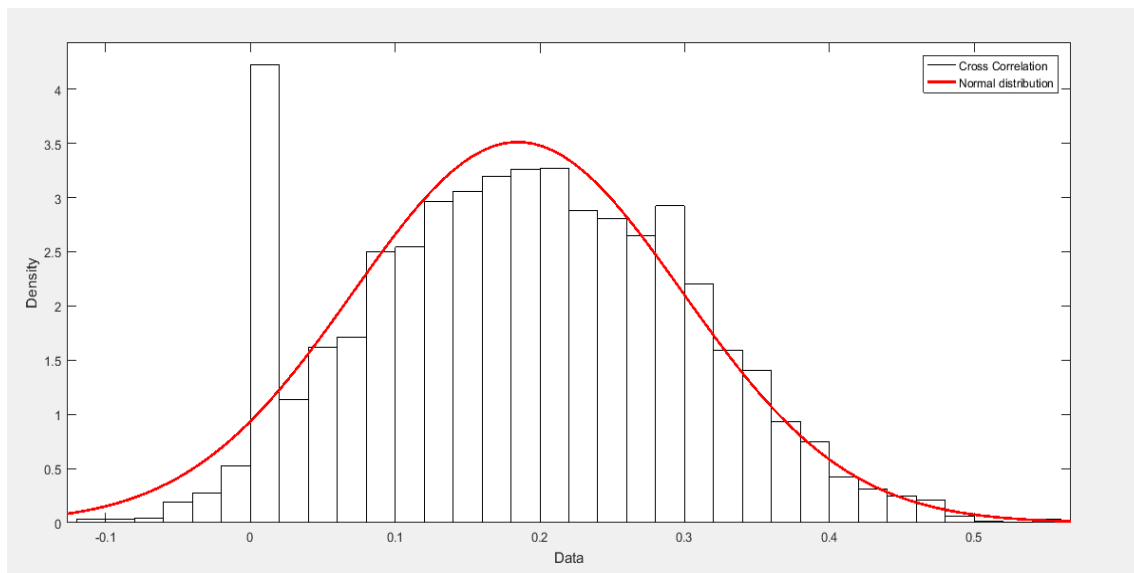
The Y axis of the histogram shows the average offer quantity at level one during April 2015. The X axis splits the continuous trading session in one-minute time intervals.

Figure 3. 6 Average level 1 bid quantity



The Y axis of the histogram shows the average bid quantity at level one during August 2015. The X axis splits the continuous trading session in one-minute time intervals.

Figure 3. 7 Cross-correlation matrix distribution



This histogram shows the empirical distribution of cross-correlation coefficients from ELP relative participation to liquidity during the 572 extreme events identified in Section 3.5.