TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

JAAKKO VALLI

PERIODICITY, CLUSTERING AND PRICE IMPACT OF LIMIT ORDER BOOK EVENTS: AN EMPIRICAL INVESTIGATION OF HELSINKI STOCK EXCHANGE EQUITIES

Master of Science Thesis

Examiner: Prof. Juho Kanniainen
Examiner and topic approved by the Faculty Council of the Faculty of Business and Built Environment
31th, July 2016

ABSTRACT

JAAKKO VALLI: Periodicity, Clustering and Price Impact of Limit Order Book Events: an Empirical Investigation of Helsinki Stock Exchange Equities
Tampere University of Technology
Master's Degree Programme in Industrial Engineering and Management
Master of Science Thesis, 69 pages
July 2016
Major: Industrial Management
Examiner: Professor Juho Kanniainen
Keywords: limit order book, price impact, periodicity, hazard-rate,clustering, algorithmic trading, high-frequency trading

This thesis studied the empirical properties of order flows in limit order books of 21 Helsinki Stock Exchange-traded stocks in a period of 700 trading days during 01.06.2010 – 26.02.2013. I examined measures of clock-time periodicity, clustering and price impact. The order book events under examination were best level limit order submissions, best level limit order cancellations and trades. I also examined how the arrivals of all submissions, cancellations and trades are affected by these events. The empirical analysis was aimed to uncover: 1) How persistent and general these properties are among different calendar time periods and companies? 2) What are the differences in these properties compared to results of the empirical literature using data from Helsinki stock exchange and around the world?

I found that 1) the events have a strong intra-second periodicity pattern, with the majority of event arrivals taking place near the start of a second. This pattern is fairly consistent among the studied companies and across the time period. Compared to the other empirical literature this effect was found to be qualitatively similar but stronger. I also showed that 2) event arrivals of all event pairs are strongly clustered beyond what could be expected based on persistent intra-day periodicity. The clustering was also found to have a persistent shape although its level differed between calendar time partitions and companies. The results on clustering were also qualitatively similar to earlier studies. Furthermore, I also found evidence that 3) the clustering pattern has a response peak at a time that varies between companies and seems to trend down (at -4,8 – -19,0% yearly) and approach 23 ms. Additionally, there are several secondary peaks at even intervals such as 1000 ms, 1500 ms and 2000 ms, which are most likely caused by order splitting. Compared to the literature the time of the first response peak is large, but I argue that this might be because higher (lower) network latency of algorithmic traders in my (other) studies data sample. I also found that 4) there is a long-term price impact, that is symmetric between sides, associated with market orders (0.019 – 0.055%) , best level limit orders (-0.0027 – -0.0071%) and, best level cancellations (0,0047 – 0,0073%) that is in line with the results of empirical studies and consistent with theory. Finally, I show that 5) the price impact measures vary between calendar time partitions and companies.

TIIVISTELMÄ

JAAKKO VALLI: Tarjouskirjatapahtumien Jaksollisuus, Klusteroituminen ja Hintavaikutus: Empiirinen Tutkimus Helsingin Pörssin Osakkeilla
Tampreen Teknillinen Yliopisto
Tuotantotalouden koulutusohjelma
Diplomityö, 69 sivua
Heinäkuu 2016
Pääaine: Teollisuustalous
Tarkastaja: Prof. Juho Kanniainen
Avainsanat: tarjouskirja, hintavaikutus, klusteroituminen, algoritminen kaupankäynti

Tämän diplomityön tavoitteena oli tarjousvirtojen empiirisiten ominaisuuksen tarkastelu ja kuvaaminen käyttäen aineistona Helsingin Pörssin 21:n Large Cap listan yhtiön 700 kaupankäyntipäivän tarjousvirroista aikavälillä 01.06.2010-26.02.2013. Aineistolle muodostettiin tarjousvirtojen ominaisuuksia, jaksollisuutta, klusteroitumista, ja hintavaikutusta, kuvaavia mittareita. Analyysissä tutkittiin mittareita kolmen ensisijaisen tapahtumatyypin: kauppojen, parhaan tason rajahintatarjousten perumisten ja parhaan tason uusien rajahintatarjousten saapumisten, seurauksia/assosiaatioita. Klusteroitumisen osalta tutkittiin lisäksi em. tapahtumien vaikutusta kolmen toissijaisen tyypin: kauppojen, kaikkien tasojen rajahintatarjousten perumisten ja kaikkien tasojen uusien rajahintatarjousten, saapumiseen. Näihin liittyen selvitettiin 1) Kuinka pysyviä ajassa ja yleisiä yhtiöiden välillä havaittavat ominaisuudet ovat? 2) Mitä eroa näillä tuloksilla verrattuna kirjallisuudessa raportoituihin sekä muilta markkinoilta, että Helsingin pörssistä tehtyihin havaintoihin?

Tärkeimpiä havaintoja olivat: 1) tarjouskirja tapahtumilla oli voimakas sekuntin sisäinen jaksollisuusrakenne jossa suuri osa tapahtumista tapahtuu sekuntin alkuosassa ja rakenne oli hyvin samankaltainen sekä yhtiöiden että ajan jaksojen välillä. Kirjallisuuteen verrattuna löytämäni jaksollisuus oli kvalitatiivisesti samankalataista mutta voimakkaampaa. Huomasin myös, että 2) tapahtumien sappuminen kaikkien tutkittujen tapahtumatyyppiparien välillä oli voimakkaasti klusteroitunutta. Kuvion muoto oli samankaltainen myös eri aikavälien ja yhtiöiden välisessä vertailussa, sekä kvalitatiivisesti samanlainen kuin aikaisemmissa tutkimuksissa. 3) klusteroitumis kuviossa oli havaittavissa ”vastepiikki”, jota edeltävä aika lyhenee tapahtumaparista riippuen keskimäärin -4,8– -19,0% per vuosi tutkimusaineiston sisällä kalenteri ajan edetessä ja vaikuttaa lähestyvän 23 ms. Lisäksi myös myöhemmillä, tasaisilla intervalleilla (1000,1500,2000 ms) esiintyy, todennäköisesti tarjousten paloittelusta johtuvia, vastepiikkejä. Kirjallisuuteen verrattuna vastepiikkiä edeltävä aika oli pitkä, mutta tämä voi selittyä algoritmikauppaa käyvien korkeammalla (matalammalla) verkkolatenssilla minun (muiden) aineistoissa. näytän lisäksi, että 4) markkinhintatarjousten (0.019 – 0.055%) ja rajahintatarjousten (-0.0027 – -0.0071%) ja tarjousten perumisten (0,0047 – 0,0073%) saapumisten pysyvä hintavaikutus oli suuruudeltaan aikaisemman tutkimuksen ja teorian kanssa linjassa. Lopulta: 5) Hintavaiktuksen suuruus vaihtelee kalenteriaikojen ja yhtiöiden välillä suuresti.

PREFACE

Thanks to my supervisor Juho Kanniainen for intense involvement, insightful comments throughout the course of this project. Also I would like to express my gratitude to all the Financial Engineering team members who I have had helpful discussions about developing the thesis. Thanks to CSC Finland for providing me with the indispensable computing resources that made the empirical analysis possible. Finally, thanks to Kiki for being understanding and supportive during the thesis writing process.

Tampere, 31th July 2016
Jaakko Valli

CONTENTS

LIST OF TERMS, ABBREVIATIONS AND SYMBOLS

Limit order book	A security, venue and currency specific queuing mechanism for limit orders.
Limit order	A pre-commitment to buy or sell a predefined quantity of security at a predefined price.
Market order	order to buy or sell a predefined quantity at the current market price.
Active limit orders	The set of limit orders that are awaiting execution in a given limit order book at a given time.
Algorithmic trading	The use of computer algorithms to aid in the submission and cancellation of orders and sometimes also making of certain trading decisions.
High-frequency trading	Subcategory of algorithmic trading, characterised by application of latest technology and exchange proximity or co-location to gain advantage by achieving the lowest possible market latencies.
Exchange co-location	Moving of ones AT servers near
Market latency	Trader and situation specific lag between something happening in the market and the trader being able to respond by e.g. cancelling a limit order. The sum of traders network latency and traders processing time.
Agency algorithm	An execution algorithm that puts to action the plan to buy or sell a given large quantity of a security by breaking it down into smaller limit and market orders that are executed over time in order to control price impact, trading costs and average execution price.
Clock-time periodicity	Statistical property of events to occur at certain times with higher probability than others as opposed to being evenly distributed.
Event clustering	Statistical property of events with tendency occur in clusters with small intervals in between events.
Price impact	The change in $b(t)$ and/or $a(t)$ associated with an event such as a market order arrival.

Direct cause clustering	Clustering of two events where the mechanism behind the clustering is that one of the events causes the other.
Common cause clustering	Clustering of two events where the mechanism behind the clustering is that both events have the same cause behind them.
LOB	Limit order book
AT	Algorithmic trading
HFT	High-frequency trading
AA	Agency algorithm
$a(t)$	Ask price at time t.
$b(t)$	Bid price at time t.
$m(t)$	Mid price at time t.
$s(t)$	Mid price at time t.
$\mathcal{L}(t)$	Set of active orders at time t.
$\Delta_x(t)$	The tick level of order x at time t.

1. INTRODUCTION

Over the last decade or so limit order books have become the predominant market structure around the world. All aspects of securities trading: price, liquidity and volatility are ultimately driven by the process order book events, i.e. voluntary submission and cancellation of limit and market orders. However, the free decision-making process with multiple actors and the emerging market dynamics has proven difficult to study. Realistic and comprehensive modelling of the process is so far missing although a lot of effort has been made towards that goal and certain stylised facts have been established (Gould et al., 2013).

At the same time, the securities trading venues have seen a development of increasing speed of information transfer and processing, resulting in an increased volume of order flow. The advantage gained through speed has driven for the adoption of new technology, and currently, a major part of order flow activity takes place in the millisecond environment; embedded in chains of interrelated order cancellations and submissions, where events follow each other with intervals of 100 ms or less (Hasbrouck and Saar, 2013). In the data sample of this thesis, algorithmic trading and high-frequency trading dominate the order flow with just HFTs share of submitted orders reaching over 80% in some stocks (Tuominen, 2012).

1.1 Research Questions and Limitations

Not many analyses of Helsinki stock exchange order book data exist to date and the few that exist focus on studying a very limited sample and/or do not address the millisecond level phenomena. This thesis aims to fill that gap by examining several non-parametric measures of order book activity in a millisecond time resolution using a fairly extensive dataset that allows the study of the persistence as well as inter-company generalizability of these measures. I study the periodicity, clustering and price impact measures and based on these answer 1) How persistent or general these results are among different calendar time periods and companies. 2) What are the differences compared to results based on empirical data from around the world and previous studies in Helsinki stock exchange.

The aim of the empirical portion of this thesis is to provide a broad, descriptive results that can be used to generate new research ideas concerning the dataset. In my analysis, I aim to find and describe the order flows properties that have been examined in various other studies: periodicity, clustering and price impact. Since AT and HFT have become so important. I focus my analysis on the very short term millisecond environment. Also, since I have a rather large sample, I can examine if the flow properties change through calendar time or between the companies.

I study the most important events, as argued by Cont et al. (2013): the trades, best level cancellations and best level submissions are all considered as triggering, primary events. Additionally, for the inter-event analysis of clustering another set of events, labelled secondary, is used. Best level submissions and cancellations are identified by the based their tick level $\Delta_x(t) = 1$ at the time t of submission (cancellation) of order x, meaning that depending on side either $p_x = b(t)$ or $p_x = a(t)$.

The analysis studies three *primary* event types namely

1. trades against bid[1] (ask[2]) side, BTR (ATR), or both together, TR;
2. best bid (ask) level cancellations, BC1 (AC1), or both together, C1; and
3. best bid (ask) level cancellations, BS1 (AS1), or both together, S1.

Additionally, there are three types of *secondary* events

1. trades against bid (ask) side, BTR (ATR), or both together, TR;
2. any bid (ask) level cancellations, BCX (ACX), or both together, CX; and
3. any bid (ask) level submissions , BSX (ASX), or both together, S1.

1.1.1 Research question

The empirical analysis is meant to allow answering the following questions:

1. What are are the differences in the (1a) periodicity, (1b) clustering and (1c) price impact results between the data partitions, with data split based on calendar time, companies and quantity quantiles?
2. What are are the differences in the (2a) periodicity, (2b) clustering and (2c) price impact results based on my data sample compared and the established empirical results in limit order book literature?

In other words, the first part of the analysis is about examining whether the results are persistent through time and general between the sample companies, try to find trends and determine which properties are stable over time. The second part is about comparing my results to the established empirical results. Furthermore, in the discussion part of the thesis, I will attempt to come up with explanations

[1] indicating an ask side market order
[2] indicating an bid side market order

for the differences and ultimately derive new research ideas based on unanswered questions.

1.1.2 Limitations

One clear limitation of this study is the choice of events and event pairs. While it can be argued that the event types that are selected are the most important and interesting ones there are still classes of events that might be interesting left out only because the line needs to be drawn somewhere. Attractive candidates to add as primary events would have been e.g. quote changes for spread decreases and increases separately and limit orders of varying degrees of aggressiveness, ranging from very aggressive inside spread limit orders to least aggressive several levels behind the best level.

Another limitation concerns the methodology and the descriptive nature of the analysis. Because the objectives are rather broad, this study is not about producing statistical tests. Although in many cases the sheer number of observations will make it obvious that observed phenomena are not caused by noise.

Additionally, there are some technical limitations in the data set itself since the time resolution is limited to one millisecond, but sometimes there can be several events occurring during that time the measures are not always correctly defined. E.g cancellation on the best level can be left out of the analysis if the tick level changes intra-millisecond. Also, the existence of hidden, automated and atypical time-in-force controlled orders and especially the lack of transparency makes it difficult to interpret some of the results.

1.2 Basic limit order book terms and notation

A limit order book (LOB) allows traders to submit and and cancel *limit orders*, public commitments to buy (on bid side) or sell (on ask side) a predefined quantity of the traded asset at a predefined price. Limit orders that are not yet executed or cancelled, the *active limit orders*, will wait to be executed in a priority queue. Lower (higher) price buy (sell) orders get higher priority; when price is tied orders that were submitted earlier get higher priority.

Additionally, Traders can submit *market orders*, orders that match at least one of the existing limit orders of the opposite side (buy or sell) resulting in an immediate execution against the highest priority active limit order(s). These three types of events: (limit order) submissions, (limit order) cancellations and market order (submissions) alter the set of active limit orders or the *order book state*.

The basic limit order book notation used in this thesis is based on the notation presented in Gould et al. (2013):

1. An order $x = (p_x, \omega_x, t_{b,x}, t_{d,x})$ submitted at time $t_{b,x}$, destroyed (executed or cancelled) at time $t_{d,x}$, with price p_x and quantity ω_x is a commitment up to quantity of $|\omega_x|$ to either buy ($\omega_x > 0$) units of the asset at price, equal or less than p_x, or sell ($\omega_x < 0$) at price equal or greater than p_x.

2. an order book event $e = (p_e, \omega_e, t_e)$ is a limit order submission, cancellation or a trade (indication of a market order submission) submitted at time t_e, with price p_e and quantity $|\omega_e|$ to either buy ($\omega_e > 0$) or sell ($\omega_e < 0$) side.

3. A LOB is defined by the traded asset, currency, venue and *resolution parameters* (σ, π). The lot size σ of given LOB is the smallest quantity allowed to be traded in the LOB (in modern equity markets this is almost always 1 and can be ignored). The tick size π is smallest possible price increment in the LOB, such that if p.

4. *Active orders* at time t are orders for which holds $t_{b,x} \leq t < t_{d,x}$. At a given time t a LOB has a state $\mathcal{L}(t)$, which is the set of all orders that are active at time t. $\mathcal{L}(t)$ can be partitioned into sets of active bid (buy) orders $\mathcal{B}(t)$, and active ask (sell) orders $\mathcal{A}(t)$, for which $\mathcal{B}(t) \cup \mathcal{A}(t) = \mathcal{L}(t)$ and $\mathcal{B}(t) \cap \mathcal{A}(t) = \varnothing$.

5. Bid price $b(t)$ is the highest price among active buy orders at time t. Ask price $a(t)$ is the lowest price among active ask orders at time t. Mid price is the average of bid and ask prices.

$$m(t) = \frac{a(t) + b(t)}{2}, \quad b(t) = \max_{x \in \mathcal{B}(t)} p_x, \quad a(t) = \min_{x \in \mathcal{A}(t)} p_x \tag{1.1}$$

6. The bid-ask spread at time t is the difference between bid and ask price $s(t) = a(t) - b(t)$.

7. For price p, the *bid-relative price* is $\delta_b(t, p) = b(t) - p$ and the *ask-relative price* $\delta_a(t, p) = p - a(t)$. For a given order x at time t the *same side relative price* $\delta_x(t)$ is the price relative to the same sides best price level. An alternative definition is the the *opposite side relative price* $\delta_x^*(t) = \delta_x(t) + s(t)$ is the price relative to the opposite sides best price level.

$$\delta_x(t) = \begin{cases} \delta_b(t, p_x) & \text{if } \omega_x > 0 \\ \delta_a(t, p_x) & \text{if } \omega_x < 0 \end{cases} \tag{1.2}$$

8. An existing orders *tick level* at time t can be defined using opposite side relative price: $\Delta_x(t) = \delta_x(t)/\pi$ in which case $\Delta_x(t) \in \mathbb{N}$. or same side the same relative

price: $\Delta_x^*(t) = \delta_x^*(t)/\pi$, and then $\Delta_x(t) \in \mathbb{Z}$.

9. The *depth* available at bid-side with price p at time t is $d_b(p,t)$ and likewise ask-side depth is $d_a(p,t)$ where

$$d_b(p,t) = \sum_{x \in \mathcal{B}(t), p_x = p} \omega_x, \quad d_a(p,t) = \sum_{x \in \mathcal{A}(t), p_x = p} \omega_x. \tag{1.3}$$

2. PRIOR RESEARCH ON PROPERTIES OF LIMIT ORDER BOOK MARKETS

In the following sections I will cover some of the most relevant publications concerning electronic limit order book and algorithmic trading related empirical findings. I concentrate on empirical results and conversely cover theoretical modelling as little as possible. I review the literature published between the year 2000-2015 with very few references before that. However, especially concerning empirical properties I emphasize studies published in 2010 or after because many of the older studies observations are nearly obsolete since algorithmic trading has increased the sheer volume of order flow so much. The first section covers research related to trading in limit order books form a traders perspective, the second section deals with algorithmic trading and high frequency trading and the third with empirical properties of limit order books concentrating on aggregated order flow and limit order book state related phenomena.

2.1 Trading in Electronic Limit Order Book

The order flow in LOB is driven by the decisions made by individual traders, so it is necessary to understand the traders motivations in order to understand the phenomena observed in order flows. Cont et al. (2012) have studied the traders' optimisation problem when they are contemplating trading in a limit order markets. After the the decision to buy or sell a certain amount of given stock has been made, the trader faces a set of options regarding the exact details of the execution. Particularly, the trader has to make choices about 1) venue; how to split the order and into parts and submit them across multiple venues, 2) scheduling; how to split the order and submit the parts over time, and 3) order type; how to choose between market and limit orders and where to place the limit orders.

Ultimately the trader faces a trade-off between two evils, namely limit orders' *execution risk*[1] and market orders' *execution cost*. Simply put, the limit orders carry risk and require patience to execute. Market orders carry little risk and execute immediately at the quoted price. However, on average the market order using buyer

[1] AKA *risk of non-execution* and *risk of adverse selection*.

(seller) ends up paying (receiving) a higher (lower) price for the traded shares compared to what is expected using a limit order. The execution cost associated with the market orders can be measured in terms of excess cost on top of the current mid price and is caused by the decision to accept the current ask (bid) price.

Copeland and Galai (1983) have argued that limit order submitters can be seen as having released free[2] options to other traders, which leads to limit orders having a disadvantage in the sense that they can be picked off by other traders who have private information, or who are able receive and process public information faster. This called adverse selection. I follows that a cancellation is a sign that the limit order submitter considers it is no longer worth the advantage because either that the value of being in the queue has diminished, or that the risk of giving out the free option has risen.

2.1.1 Information Arrival and Processing

Because the risk and reward associated with a limit order may change, it is beneficial to assess them frequently. Then the trader can adjust the execution strategy (venue, scheduling, submission) and the submitted orders as necessary. Otherwise changing conditions can make the submitted orders go stale.

The traders risk associated with a limit order are defined by 1) speed information arrival, 2) traders relative speed of information processing and taking action. The source of the information can be either exogenous such as news or stock exchange releases, or endogenous, directly related to the order book (Bouchaud et al., 2002; Hautsch and Huang, 2012). The relative speed in information processing and taking action is a trader-specific attribute that determines how quickly, relative to other traders, the trader can access, process and act on new information. Gomber et al. (2011) highlight the importance of considering *relative* speed asserting that execution risk is increased in traders with low relative speed.

Information arrival speed that is something that will change from moment to moment. Siikanen et al. (2016) show that the traders account for this, and the order book state goes through certain predictable changes during the anticipation, release and immediate aftermath of stock exchange releases[3]. These effects can be seen as a direct consequence of the traders adjusting their strategies in response to expected increase in information arrival speed. On the other hand Linnainmaa (2010) shows

[2]Free in the sense that they receive no payment. Because of time priority they do gain the advantage of holding a higher place in the priority queue compared orders submitted to the same price level at a later time.

[3]They show in particular that this applies to markets I study in the empirical analysis although the references therein show that this is the case in other markets aswell.

that failing to account for the increased information arrival speed by cancelling limit orders prior to announcement causes significant losses to household investors. Importantly this is because they lack the ability to monitor and readjust their orders and thus fall victim to adverse selection, which is a concrete example of a disadvantage of relative speed and how it is conditional to the exogenous information arrival speed.

Information endogenous to order flow has equal results in the literature. Bouchaud et al. (2002); Hall and Hautsch (2004); Hautsch and Huang (2012) have shown evidence that traders consider order flow events: trades, cancellations, and submissions, as sources of new information too, because it reveals intentions of other traders. The other side of the coin is, as several authors (Cont et al., 2013; Eisler et al., 2012; Hautsch and Huang, 2012) argue, that when a trader wants to sell or buy a large quantity in a LOB they should consider the effects their actions will have on the market. These effects are referred to as the *price impact*: the change in bid and ask prices associated with order book events, and*market impact* : the set of all effects in the entire order book state. Bouchaud et al. (2008) further divides these into two subcategories: *immediate impact* that takes place the instant an event takes place, and *permanent impact*. Additionally, a related concept, the *resiliency* of order book that is, as defined by Degryse et al. (2005), the "speed of the recovery of the market (in terms of price, depth and spreads) after a relatively large shock defined as trade that increases the bid ask spread". Their definition of resiliency reflects the idea that immediate impact may be larger than the permanent impact, and this is what they also empirically confirm.

2.2 Algorithmic Trading and High-Frequency Trading

Hendershott et al. (2011) defines *algorithmic trading* (AT) as "... the use of computer algorithms to automatically make certain trading decisions, submit orders, and manage those orders after submission.". According to Chaboud et al. (2014) the two main features of AT in the literature are: 1) The AT systems' speed advantage over humans, 2) potentially high correlation of algorithmic traders strategies and actions. *High-frequency trading* (HFT) is defined as a subset of AT distinctive most importantly by a low latency, but also several other properties.

Currently, most of the typical LOB trading venue's orders originate from AT systems and also ATs share of trading volume is high: according to (Gomber et al., 2011), in 2010 (the start of the period my empirical data is from) trade volume were between 19% and 70% for European and US marketplaces. Characteristics that apply to both of algorithmic and high-frequency trading are: "1) pre-defined trading decisions, 2) use by professional traders, 3) observing market data in real time, 4) automated order submission, 5) automated order management, 6) no human intervention, and 7) use of direct market access".

I adopt the common[4] classification of AT into two subcategories: *agency algorithms* and HFT. This division is also consistent with how Gomber et al. (2011) describe HFT and non-HFT AT. According to them properties of AT which do not apply to HFT (I use the term agency algorithms to describe this group): "1) Used in agent trading, 2) object to minimize market impact for large orders, 3) goal to achieve a particular benchmark, 4) holding periods from days to months, and 5) object to work an order through time and across markets". Similarly they list several properties of AT that apply exclusively to HFT: "1) Very high number of orders 2) rapid order cancellation, 3) proprietary trading, 4) profit from buying and selling (as middleman), 5) no significant position at end of day (flat position), 6) very short holding periods, 7) extracting very low margins per trade, 8) low latency requirement, 9) use of colocation/proximity and individual data feeds and 10) focus on high liquid instruments." They also list a set of other structural developments that they argue AT and HFT has gone hand in hand with: "1) new market access models and fee structures, 2) significant reduction of latency and 3) the fragmentation of order flow". The listings illustrate that the reasoning behind the success of AT and HFT is that they solve the problems related to 1) controlling unfavourable market impact (agency algorithms) 2) and fast information processing (HFT).

As defined by Hasbrouck and Saar (2013) *market latency*[5] is the sum of network latency and processing time. Network latency is the time delay that is caused by the physical distance from the and it is calculated by multiplying the duration of a one-way trip, which is based on factors such as distance, with 2. First, the latest information on current market state needs to reach the decision maker, and then there is a second delay while the response makes it's way back to the exchange. Processing time is the time taken until the next action can be sent to the exchange, i.e. it is the time that it takes for the AT system to decide the on the next action. According to Gomber et al. (2011) reducing market latency is crucial because latency inhibits reacting to changing conditions and changes that occur in the market conditions during the latency lag are sources of additional risk.

2.2.1 Agency Algorithm Trading Strategies

Agency algorithms (AA) serve primarily the execution need of, such as pension funds, and though they come in many flavours, all of them are aimed at solving the

[4] It is also utilised by Hasbrouck and Saar (2013) and NASDAQ (2016)

[5] They talk about just latency, but since there is more than one kind of latency it is useful to be explicit

same problem. Namely, how to grow or decrease a position in a given instrument by a relatively *large total amount*, while avoiding a costly *market/price impact*, achieving an execution at a *predefined price*[6]. Generally, this involves breaking the execution into several small orders, but the differences arise from how exactly this is done.

Participation rate algorithms (PRAs) are a subcategory of agency algorithms that participating in the market using a preset trade quantity. By limiting their participation, the algorithms aim to hide the trading intention to avoid having the prices move against them. VWAP algorithms will set the quantity target relative to realised trading volume in the order book. As their name suggest, they will get an average price close to the *volume weighted average price* during the execution period because they match the volume traded in the order book. TWAP algorithms aim to do the same except they spread the traded quantity roughly evenly based on time and achieve the *time weighted average price Implementation shortfall algorithms* (ISAs), combine a historical and real-time data in order to minimize the market impact of a large order by finding the optimal execution strategy in respect to order types, timing and venue. (Gomber et al., 2011)

All of the agency algorithms success is based on their ability to conceal the trading intention. AA activity creates a front-running opportunity for someone who can detect their intentions from the stream of orders. Since some of the actions of high-frequency trades are claimed to "impair the prices" for other market participants such as the large buy-side institutions, these HFTs get called "predatory", which highlights the dynamics of conflicting interest that exist between these two categories (Gomber et al., 2011). Arnuk and Saluzzi (2009) describe one such HFT strategy.

2.2.2 High Frequency Trading Strategies

Gomber et al. (2011) state that HFT is not a trading strategy in itself, just the application of technology and colocation and in fact HFT systems are applied to traditional trading strategies to boost their returns. According to Chlistalla et al. (2011) HFT systems are operated exclusively on proprietary terms and the operators use their own capital to execute the strategies. They list three main subcategories of strategies that HFTs engage in: 1) liquidity provision 2) statistical arbitrage and 3) liquidity detection. This section gives a brief description of the categories highlighting the HFT-specific aspects.

[6]not predefined in an absolute sense but rather relative to what happens in the market at the time the execution takes place

Liquidity provision is the most common HFT strategy. E.g.Hagströmer and Norden (2013) find that HFT market making amounts to 63-72% of total HFT trading volume and 81-86% of total HFT limit order submissions of OMX Nasdaq Stockholm in 2011-2012[7]. HFT liquidity providers will often operate similarly to formal market makers, but lack the obligations usually imposed on them. They have revenues from two sources, 1) earning the bid-ask spread as they provide liquidity, 2) incentives from the trading venues that grant rebates and reduced fees in return for the increased market quality and attractiveness. (Chlistalla et al., 2011; Gomber et al., 2011) HFT technology is leveraged in form of faster adjustment of quotes e.g. when the algorithm thinks that it might be in danger of being picked off by informed traders.

Statistical arbitrage is a traditional strategy that fits particularly well together with the advantages of HFT. HFTs engage in arbitrage trading in similar ways to the traditional arbitrageurs, but outperform them by leveraging the latest technology to achieve the lowest possible latencies. Because benefiting from arbitrage is about reacting to observed discrepancies, the HFT arbitrageurs operate primarily as liquidity takers. (Gomber et al., 2011)

Liquidity detection is a practice aimed at gathering information about the hidden liquidity in a given order book. This is done by sending out small orders to the order book (AKA ”pinging”) to see if they match. The aim of this is to reveal hidden liquidity and gain an advantage by acquiring a fuller set of order book state information compared to other market participants and thus being able to make better predictions about how the prices are likely to evolve. Since this strategy applied to uncover how other algorithms will respond to changes in the order book state, the practice is often referred to as ”sniffing out” other algorithms. (Gomber et al., 2011)

For a description of HFT activity in practice, one can look at one of the methods used to detect it. Hasbrouck and Saar (2013) take up the challenge of detecting HFT[8] activity from unlabelled order flow data and show that their measure, constructed to count the number of active ”strategic runs”, is highly correlated with the actual HFT activity. Their definition of a strategic run, in short, is a sequence of fast ($\leq$ 100 ms) limit order submissions and cancellations (and possible resubmissions). They show these runs often consist for several hundred resubmissions and interact with each other by having multiple parties seemingly locked in a feedback loop of actions and responding actions. The strategic runs possibly constitute the HFTs following liquidity provision and detection oriented strategies, but because of the

[7] overlapping time with my dataset
[8] they use the term low latency AT

orientation of the measure towards limit orders, it will probably not capture the activity of arbitrageurs as they use market orders.

2.2.3 Impact of AT and HFT on the Market

Hendershott et al. (2011) study NYSE stocks AT order flows in 2003 and finds AT to be narrowing the spreads, reducing adverse selection, and reducing trade-related price discovery, especially in the case of large stocks. Overall they conclude that AT improves both liquidity and the informativeness of the bid-ask prices. Chaboud et al. (2013) conduct an empirical investigation of the FOREX market of 2003-2007 and conclude that the rising involvement of ATs during that period has reduced the availability of triangular arbitrage opportunities and reduced autocorrelation of high-frequency returns. They also find evidence that AT systems operate on a less diverse set of strategies and that it may increase volatility intermittently. Brogaard et al. (2014) investigate HFT order flows in 2008-2009 dataset from NASDAQ (US) and find several effects 1) HFTs are likely to follow a price reversal strategies based on order imbalances, 2)in the US HFTs' annual revenue is about 3 billion USD, 3) HFTs do not systematically front run non-HFTs, 4) HFTs engage in a more narrow set of strategies compared to non-HFTs, 5) HFTs engagement in trading is not very sensitive to volatility increases (i.e. there is no evidence of systematic mass-withdrawal), 6) HFTs provide a large contribution to the price discovery process, 7) HFTs provide the best quotes for a significant portion of the day, but at a their quotes are only about one-fourth of the book depth compared non-HFTs, and 8) There is no evidence that HFTs increase volatility, but they may in fact instead reduce it.

Gould et al. (2013) review the empirical literature related to effects of HFT and conclude that there is somewhat conflicting evidence. On one hand, several authors (Hendershott et al., 2011; Chaboud et al., 2013; Brogaard et al., 2014; Gerig, 2015) have reported primarily positive effects such as that HFT decreases adverse selection, increases market quality by reducing spreads, increasing informativeness of quotes, increasing liquidity, and increasing market stability. On the other hand Biais and Woolley (2011); Kirilenko et al. (2011) have taken a more critical stance and argued that HFT exploits an unfair speed advantage, HFT activity, in fact, decreases liquidity and is prone to increase volatility especially in situations of market turmoil, when it is most needed. Gomber et al. (2011) note the same negative arguments but find a much larger volume of literature arguing that HFT has primarily positive effects on market quality.

Chlistalla et al. (2011) conclude in their review that there is some evidence of positive effects and no evidence on negative effects, but they list some concerns and unresolved issues that are primarily related to HFT liquidity provision: 1) HFTs don't have an obligation to post quotes and thus it is conceivable that their liquidity may disappear under volatile market conditions when it is most needed. 2) HFTs post a lot of small orders and do not contribute much to the total market depth thus their orders may cause that large orders execute against a lot of small orders. 3) HFT quotes are barely accessible because the orders they post are very short lived.

One possible explanation to reconcile the somewhat conflicting results, that is discussed by a few authors (Biais and Woolley, 2011; Brogaard et al., 2014) is that HFTs cause liquidity suppliers to suffer adverse selection costs, and this could, especially in times of turmoil lead to sources of non-HFT liquidity to pull out. So while HFTs participation might have positive direct consequences on measures of market quality there still might, especially under certain special conditions, be a net negative effect due to the indirect effect of scaring away liquidity.

2.3 Empirical observations in LOB Markets

Several statistical regularities exist for a large enough portion of the markets so that they may be called *stylized facts* of the limit order book markets. This section deals with the empirical observations that are specific to limit order book markets and.

2.3.1 Event Conditionality on the Limit Order Book State

Several studies have shown that current order flows can be explained, in a large part, as dependent on the current order book state (Ellul et al., 2003; Hall and Hautsch, 2004; Ranaldo, 2004; Hollifield et al., 2004; Lo and Sapp, 2010). Studying these dependencies is quite difficult because 1) It is tricky to condition order flows on the vast state space of limit order books. (Parlour and Seppi, 2008). 2) When there are fast changes in the state latency of limit order books makes it unclear what is the state that a particular event should be conditioned on. (Gould et al., 2013)

The most prominent findings are that: 1) Limit orders with larger relative prices are associated with lower quantities(Bouchaud et al., 2002; Maslov and Mills, 2001) 2) Wider (more narrow) spreads are followed by higher probability of limit (market) order arrivals, and larger quoted depth causes competition to supply liquidity. I.e traders probability to resort to market orders is negatively correlated with their cost relative to limit orders and limit order submission probability is negatively correlated with the depth in front of the order (because it is connected to expected waiting time). The traders will want to move their orders to the front of the execution queue, even if it means they have to post more aggressive limit orders provided that the queue is long enough. (Ellul et al., 2003; Hall and Hautsch, 2004; Cao et al., 2008).

2.3.2 Event Clustering

Order flow event clustering is observed/studied in a number of empirical studies (Ellul et al., 2003; Degryse et al., 2005; Hall and Hautsch, 2004; Hasbrouck and Saar, 2013). The analysis of Australian stock exchange by Hall and Hautsch (2004) determines that, in addition, exogenous factors drive the order flow, by intermittently increasing the level of overall order flow activity, making the event arrivals highly clustered. I.e from a traders perspective: once a decision to trade has been made the optimal order submission is largely dependent on order book state and how it is expected to change, but the decision is also conditional on exogenous information arrival. Also, the events themselves cause a feedback loop triggering new adjustments by other traders. This means that there are, by my own definition, two distinct reasons for observed clustering of order flow events: 1) *Common cause clustering* between events emerges as a result of the events both being caused by the same original reason/event, while 2) *direct cause clustering* between two events is a result of one type of event causing the other type. Since it is impossible to correctly label the observed clustering to these classes, order flow models and empirical tests tend to assume implicitly that common cause clustering does not exist, i.e., that there are no unknown exogenous drivers of the process.

Degryse et al. (2005) use data from Paris Bourse to examine the order flow around different types of orders, categorised by their price relative to current best bid and ask price. They have six categories per side with the most aggressive being a market order that uses up liquidity from several levels on the opposite side of the book and least aggressive being a limit order beyond the current best level. They find 1) that buy (sell) orders are more likely to be followed by new buy (sell) orders[9], 2) That orders of certain side and level of aggressiveness will make the same kind of orders more likely to occur immediately after 3) that aggressive (market) orders tend to expand the spread 4) persistence in the order flow patterns, i.e., a given type of events are responded similarly in samples from different time periods. 5) Aggressive orders tend to take place when the liquidity becomes exceptionally good.

Limit order book models of Zhao (2010) and Toke (2011) are build on the assumption that events are clustered, and their models have the "self-exciting" property. Meaning that event arrival increases the arrival rate of future events briefly. This kind of models produce the kind of event clustering effect that is present in the real data, but the problem is that they attribute all clustering to the self-exciting mechanic and ignore the effect of exogenous information, which conflicts with the findings of Hall and Hautsch (2004)

More recent evidence on the reasons of clustering comes from Toth et al. (2015) studying the persistence of order flows using data from London Stock Exchange with identifiers of brokers. They state that the two possible explanations for the market order sign autocorrelation (more buys follow buys and sells follow sells) are 1) herding behaviour (positive correlation between agents decisions) and 2) order splitting (positive correlation within single agents sequence of actions). They conclude that in their sample in less than a few hours time span order splitting rather than herding is the reason behind the observed sign autocorrelation.

[9] Ellul et al. (2003) also concludes that positive (negative) returns produce more buy (sell) orders, which may be related to event clustering.

2.3.3 Order flow periodicity

Hasbrouck and Saar (2013) in their examination of NASDAQ (US) data from 2007-2008 find a distinct, persistent order flow periodicity pattern. Which, as they argue, emerges likely from AAs' programming that contains regular patterns (e.g., specific time within the second when to evaluate the market state). The patterns do not disappear because other market participants' responding orders feed into them, rather than correct them. This explanation would make clock-time periodicity a source of common cause clustering, where the common cause is the time of day.

The intra-second periodicity pattern has to be algorithm driven also because humans cannot plausibly regularly achieve precision that would cluster the events initiated by them. On the other hand, there are clear periodic patterns in (e.g., intra-day) order flow that humans do manually get involved in. Thus periodicity itself is not new, but millisecond environment periodicity is a relatively new phenomenon as it is driven by the increased AT activity.

2.3.4 Market- and Price Impact of Limit Order Book Events

Eisler et al. (2012) list several stylised facts related to trades that they consider to be agreed upon in the literature: 1) Buying (selling) using market orders in a LOB creates an upward (downward) price impact in both $a(t)$ and $b(t)$. 2) The price impact as a function of market order quantity is concave, which means that orders with large quantity cause a price impact that is only slightly bigger than small quantity orders. 3) There is a strong autocorrelation of market order (trade) sign. However, the $m(t)$ movement is almost purely diffusive. They also identify two distinct parts of the price impact (after any event). The *direct price impact* part is the immediate price change caused by the event. E.g. A large buy order will cause an immediate jump up just by expanding the liquidity on the best level(s). The *induced price impact* part, or dynamic part is based on the change in the future event rates and their associated gaps. E.g. following the original example the induced effect would mean that new limit orders would be likely to appear on the bid side inside quote and the spread would narrow. These effects are still a part of the market orders price impact because without the original market order these new limit orders would not have appeared there. Eisler et al. (2012) framework assumes that the observed clustering is exclusively event induced (direct cause clustering.) Similar to the self-exciting models (Zhao, 2010; Toke, 2011) they neglect the possibility of effects of common cause clustering.

Hautsch and Huang (2012) conduct an empirical examination of the market impact of limit- and market orders in what is perhaps the most extensive analysis of limit order price impact dynamics so far. They find evidence that limit orders have significant long-term effects and that there is a cointegration relationship between ask and bid levels with corresponding depths. They show that limit orders have a market impact that is smaller (by an order of magnitude) than the impact of trades of similar size but none the less significant. Additionally, they show that the least aggressive limit orders placed deeper into the book have an impact, but it is smaller than that of the best level submissions. However the fact that limit orders (also deeper in the book) have a market impact leads to order book states that are less informative about the actual willingness to trade. Toth et al. (2011) argues that the collection of available limit orders would not reflect entire set of the prices at which the traders would be willing to trade because traders will often benefit from hiding their intentions until the last possible time.

Cont et al. (2013) develop the *order flow imbalance* to describe the price impact of order book events which is based on cumulates the order flow on the best levels taking into account new limit orders' volume as positive and cancelled volume or traded volume as negative flow and then summing these together. They show that: 1) this rather simple approach explains the price changes rather well. 2) The sensitivity to price changes is negatively correlated with average market depth. Their model does not even attempt to take into account the induced effects of events or any event clustering, but it is just estimating the total direct (as defined by Eisler et al. (2012)) price impacts of the arrived events.

The empirical investigation of Chinese stock market in 2003 by Zhou (2012) finds evidence that 1) there is a nearly perfect symmetry between buying and selling market orders price impact dynamics, 2) the market orders price impact can be explained by both quantity and order aggressiveness together, and 3) there are separate price impact curves for market orders that are executed partially and ones that are executed completely, as the partial executions' price impacts are not sensitive to order quantity.

2.4 Prior Research Utilizing Overlapping Data

Tuominen (2012) studies a one week period (22-26.11.2010 of 5 Helsinki Stock Exchange based, large cap companies' order book data. Using data consisting of individual accounts behind the order flows, he makes several unexpected and interesting findings that probably apply to most of my data as well: 1) most of the orders are generated by a few accounts: A single account is responsible for about 31% of all orders and about 70% orders are submitted by just five of the most active accounts and over 80% by the 10 most active accounts. 2) Among the five heavily traded large cap the HFTs are responsible for 43-82% of all orders. 3) HFT activity decreases during the trading day in each stock separately which may be an indication that the HFTs are engaging in liquidity provision strategies. 5) HFTs update their orders very quickly, and as a result, they (individually) frequently send hundreds of consecutive messages without anything being submitted in between by any other account. 6) Taking all of the events into account there is a response peak at 26 ms which he identifies as direct cause clustering. He also speculates whether the results are biassed because of periodicity or external factors like in common cause clustering. 7) The HFTs activity have short term impact on the price, with limit order book events predicting price impacts to up to 150 ms to the future.

Toivonen (2013) study two months (03.2010,09.2012) with an overlapping set of companies and find that 1) The execution rate is small for both months (8.0%,5.7%) 2) Number of limit orders submitted increase by 68% to an average of 60 thousand per day. 3) The majority of limit order lifetimes are less than 10 seconds, but excluding all but the most active stocks the majority of lifetimes fall below 2 seconds. 4) High-frequency price changes can be explained using the order flow imbalance measure, which explains an average of 57% of the mid-quote price impact. The impact has an intra-day periodic pattern with price impact at its largest in the morning. 5) The order flow imbalance is superior to the trade impact in explaining price impact. 6) Order flow impact goodness of fit increase between the two periods.

Eskelinen (2015) study the period of 2010-2012[10] using an overlapping set of companies to find that: 1) The OFI model works as expected based on the original results, with R^2 between 0.40-0.57 and diurnal effect of decreasing β(Similiar results to Toivonen (2013)). 2) There is significant clustering as measured by MM and LL-effect, showing that previous 5 minute periods number of orders (trades) predicts next period's number of trades 3) There are intra-day half hour bin periodicity patterns of event arrival rates. 4) There is a hump in order inter-arrival rate "at around 20 ms" which he hypothesises to be caused by direct cause clustering.

[10]This is almost the same time period as my empirical sample.

3. EMPIRICAL ANALYSIS

This chapter describes the empirical part of this study. The first section (3.1) gives a description of the used data set, the second section (1.1) introduces the research questions, the third section (3.2) descibes the used methods in detail, and the final, fourth section (3.3) presents the results of the empirical analysis.

3.1 Order flow data

The used data is the full order flow, based on Nasdaq TotalView ITCH feed dataset. I examine the order flows of 21 Finnish companies large cap[1] listed primary[2] stock series, during the period 1.6.2010-06-26.2.2013 consisting of 700 trading days. In the analysis the data is distributed company and calendar time wise into 21 * 7 = 147 data partitions that each contain 100 trading days of a one companies data.

List of calendar time partitions				
Index	Start date	End date	Calendar days	Trading days
1	1.6.2010	18.10.2010	139	100
2	19.10.2010	9.3.2011	141	100
3	10.3.2011	2.8.2011	145	100
4	3.8.2011	21.12.2011	140	100
5	22.12.2011	14.5.2012	144	100
6	15.5.2012	2.10.2012	140	100
7	3.10.2012	26.2.2013	146	100

Table 1: The table lists the examined calendar time partitions. Start date, end date and the number of calendar days for each date partition are given. The indices can be used to identify the calendar time partitions in the results. The number of trading days for each partition is 100.

The data set is described in tables 1 and 2 which lists the companies and calendar time partitions that are studied. These tables can also be used as a reference to

[1]Each stock series has been on the OMX Helsinki Large Cap list during the period. I have filtered out Large Cap companies that have significant trading in other currencies, e.g. Nokia. Because of the order flow fragmentation, including them would have created and "apples to oranges" comparison in analysis of company-wise differences.

[2]selected based on having the highest trading volume within the sample

List of stocks			
Index	Stock name	Isin code	BIC category
1	Kesko Oyj B	FI0009000202	Consumer Services
2	Tieto Oyj	FI0009000277	Technology
3	Outokumpu Oyj	FI0009002422	Basic Materials
4	Sampo A	FI0009003305	Financials
5	Wärtsilä Oyj Abp	FI0009003727	Industrials
6	Kemira Oyj	FI0009004824	Basic Materials
7	Nokian Renkaat Oyj	FI0009005318	Consumer Goods
8	Konecranes Oyj	FI0009005870	Industrials
9	Stora Enso R	FI0009005961	Basic Materials
10	UPM-Kymmene Oyj	FI0009005987	Basic Materials
11	Fortum Oyj	FI0009007132	Utilities
12	Sanoma Oyj	FI0009007694	Consumer Services
13	Metso Oyj	FI0009007835	Industrials
14	Elisa Oyj	FI0009007884	Telecommunications
15	Neste Corporation	FI0009013296	Oil & Gas
16	KONE Oyj	FI0009013403	Industrials
17	Cargotec Oyj	FI0009013429	Industrials
18	Orion B	FI0009014377	Health Care
19	Outotec Oyj	FI0009014575	Industrials
20	YITOyj	FI0009800643	Industrials
21	Rautaruukki	FI0009003552	Industrials

Table 2: The table lists the studied stock series giving the index, name, isin code and BIC category of each stock/company. The indices can be used to identify the companies in the results.

interpret the results as they contain each company's and calendar time partition's reference index number.

There are some clear advantages in using this particular data set in order book / order flow research. First, most of the studied companies are exclusively traded in Euros and in their respective order books hosted at OMX Helsinki, which avoids the challenges associated with liquidity fragmentation[3]. Additionally, the stocks are relatively liquid but not so liquid that liquidity effects would be unobservable (e.g. the spread would almost exclusively remain at one tick).

3.1.1 Nasdaq OMX Helsinki Market

Helsinki stock exchange has continuous trading between 09:00 and 17:25 CET. I want my results to reflect typical market conditions. Hence I cut an additional 25 minutes from the end and 30 minutes from the beginning and end up with 7.5 hours

[3]According to Gould et al. (2013) liquidity fragmentation is one of the challenges in studying LOBs

of continuous trading time during 09:30-17:00. I filter the events outside of this intra-day time interval[4].

Time in Force Attributes

In a simplified model of a LOB orders are considered to be active until traded or explicitly cancelled (i.e. order is *good-till-cancelled*). In practice, however, the exchange offers several additional time-in-force rules that can be used to exert automated control over the lifetime of an order at the time of submission. These additional rules are: *Immediate-or-cancel* (IOC), *good-till-close* (GTC) and *good-till-time* (GTT). Especially IOC can cause confusion if one assumes that traders cancel their order using a separate cancellation message during the same millisecond when in fact the process is automated and the decision pre-meditated.

Special Order Types

In addition to a standard limit order that exists until cancelled, there are a few special order types that the are allowed by the exchange rules these are Iceberg, pegged and hidden orders NASDAQ (2010).

Iceberg orders are limit orders that have been split into a visible and hidden part of quantity. As the visible part functions as a standard order and once it gets executed a new order with a new time priority is created from automatically from the hidden quantity reserve. The visible part is always published in the market feed as any other order.

Pegged orders are orders that automatically move so that they maintain a certain predefined relative price distance to $b(t)$, $a(t)$ or $m(t)$. As the best quotes change pegged orders automatically cancelled and resubmitted and these events are published in the feed. The above means that both iceberg and pegged orders updates can be a source of automated clustering. If neglected they may also give a distorted view of their submitters' market latency because if they did not exist traders could track their orders and update them just the same, except they might have a higher latency between the updates.

Hidden orders are orders whose liquidity is entirely hidden and only is discovered only when another trader posts a limit order that leads to a trade with the hidden order. These create a problem because while there is hidden liquidity between $b(t)$ and $a(t)$ those prices are not reached and

[4]Naturally I do consider all of the orders when I construct the order book states that are necessary e.g., for determining which events occur at best price levels levels.

3.2 Methods

This section introduces the methods that are used to calculate the measures and in the empirical analysis and discusses how the aggregation is conducted over the data partitions. The empirical analysis examines and produces results for the 147 data partitions, 3 primary event types, 3 secondary event types, and the 9 event type pairs using 4 basic methods:

1. *Clock-time periodicity* to measure periodicity specific to each data partition and primary event in 3 different analysis setups (covered in 3.2.1).
2. *Inter-event hazard rate* to measure inter-event clustering and event conditional periodicity specific to each data partition, event type pair and 4 side combinations of the event pair (ask-ask,ask-bid,bid-ask,bid-bid) (covered in 3.2.2).
3. *Order lifetime hazard rate* to measure order lifetime clustering specific to each data partition (covered in 3.2.2).
4. *Clock-time price impact* to measure price impact specific to each data partition, primary event, and also partitioned to 6-Quantile bins based on the primary event volume (covered in 3.2.3).

3.2.1 Periodicity measures

Millisecond-level periodicity patterns are an interesting feature of the modern markets. Hasbrouck and Saar (2013) provide evidence for persistent clock-time periodicity[5] patterns that they argue are an indication of the activity of AT systems that make periodic checks to the market state and adjust their orders accordingly. I use the data partitions that allow studying the inter-company and calendar-time-specific differences.

I measure empirical the periodicities of using the exact the method used in Hasbrouck and Saar (2013), but present the result relative to uniform distribution: For each bin edge $xs_P \in \{0, s_P, 2s_P..., \tau_P\}$[6] (where s_P is bin width and τ_P is the period window length) milliseconds the number of events where the event time $t_e(i, c, d)$ satisfies $(x-1)s_P \leq \text{mod}(t_e(i, c, d), \tau_P) < (x)s_P$ is calculated and compared to the result expected[7] (τ_P/s_P) given that events are uniformly distributed:

[5]Later on I use simply "periodicity" to refer to "clock-time periodicity".

[6]Note that the bin edge index x has nothing to do with previous definition of x referring to order.

[7]However, the τ_P becomes $\max(\tau_P, 1000 \times 60 \times 60 \times 7.5)$ because it needs to be capped to milliseconds in examined intra-day trading hours, which comes to effect in the case of the intra-day periodicity pattern.

$$\hat{P}_e(x,c,d) = \frac{\frac{1}{n_e(c,d)} \sum_{i=1}^{n_e(c,d)} \mathbb{1}_{(x-1) \leq \frac{\mathrm{mod}(t_e(i,c,d),\tau_P)}{s_P} < j}}{\max(\tau_P, 1000 \times 60 \times 60 \times 7.5)/s_P}, \tag{3.1}$$

Clock time periodicity is calculated for each 147 data partitions and for each primary type for 3 different setups $(\tau_P, s_P) \in \{(1000, 1), (10000, 10), (24\times60\times60\times10000, 60\times 1000)\}$[8]

3.2.2 Hazard rate measures

Hazard rates are chosen as the examined measure because the convenient interpretations they have: A flat curve represents a memoryless Poissonian process, while a decreasing hazard rate (DHR) curve indicates a clustered behaviour and finally increasing hazard rate (IHR) curve indicates a system where short intervals are rarer than with an independent process. Based on earlier empirical work (Hall and Hautsch, 2004; Hasbrouck and Saar, 2013) I expect to find a decreasing hazard rate curve, but the interesting question is whether there are local inclines. Since there is no parametrization or loss of data resolution involved; any shape of the function is possible and quick jumps can be observed in the millisecond resolution. Event clustering of events also explains the *induced* (as defined by Eisler et al. (2012)) part of price impact. Thus these two measures can be interpreted together to gain further understanding.

Inter-event times

The *inter-event hazard rate* is measured using a method similar to Hasbrouck and Saar (2013). The inter-event times calculation is exactly the same: For company c in time partition d, I measure the shortest time difference between $n_{e_1}(c,d)$ occurrences of e_1 events and $n_{e_2}(c,d)$ occurrences of e_2 events so that I get $n_{e_1}(c,d)$ time differences:

$$t_{e_1,e_2}(i,c,d) = \min_j(t_{e_2}(j,c,d) - t_{e_1}(i,c,d)), t_{e_1}(i,c,d) < t_{e_2}(j,c,d) \tag{3.2}$$

where $i \in \{1, 2, ..., n_{e_1}(c,d)\}, j \in \{1, 2, ..., n_{e_2}(c,d)\}$.

[8]The first setup matches that of Hasbrouck and Saar (2013).

Hazard rate calculation

The main difference between the analysis of Hasbrouck and Saar (2013) and mine is that, when calculating the hazard rates, they treat all other events as exogenous censoring events, which is an unbiased specification *only* if the censoring times are not conditional on the measured event times. In the data used in this thesis this is clearly not the case, and so I do not use censoring[9]. Calculating the hazard rate from the time differences is based on the following: Given that random variable $X \in \mathbb{R}$ has CDF $F(x) = \mathrm{P}(X \leq x)$ and PDF $f(x) = F'(x)$. Then the hazard rate function (HRF) is

$$h(x) = \frac{f(x)}{1 - F(x)}. \tag{3.3}$$

Assume there is data: n observations $t(i) \in \mathbb{N}, i \in \{1, 2, ..., n\}$ and $\mathbb{1}_{t(i) \leq x}$ is an indicator that the observation $t(i)$ less or equal to x, then empirical hazard rate function (EHR) $\hat{h}(x)$ an be calculated combining:

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{t(i) \leq x}, \tag{3.4}$$

$$\hat{f}(x) = \hat{F}(x) - \hat{F}(x-1) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{t(i) = x}, \tag{3.5}$$

and

$$\hat{h}(x) = \frac{\hat{f}(x)}{1 - \hat{F}(x)}, \tag{3.6}$$

to get

$$\hat{h}(x) = \frac{\frac{1}{n} \sum\limits_{i=1}^{n} \mathbb{1}_{t(i)=x}}{1 - (\frac{1}{n} \sum\limits_{i=1}^{n} \mathbb{1}_{t(i) \leq x})}. \tag{3.7}$$

In the emprical analysis the *empirical hazard rate function*, $\mathrm{EHR}_{e_1,e_2}(x, c, d)$, of inter-event times $t_{e_1,e_2}(i, c, d)$ for company c, and calendar time partition d is calculated up to window length τ_{HR} ms ($x \in \{1, 2, ..., \tau_{HR}\}$) by substituting $\hat{h}(j) = \mathrm{EHR}_{e_1,e_2}(x, c, d), n = n_{e_1}(c, d), t(i) = t_{e_1,e_2}(i, c, d)$ in equation 3.7.

[9]This choice is not entirely without problems: There can be bias due to confounding events. However, this choice has the benefit of enabling the estimation of the hazard rates further away from the original event.

Order life time

One special case of EHR that I examine is the *empirical order lifetime hazard rate*, (for traded orders $\mathrm{EHR}_{LTT}(x, c, d)$, and for cancelled orders $\mathrm{EHR}_{LTC}(x, c, d)$). The inter-event time observations are calculated as the time from a specific orders' creation (submission or alteration of an existing order) those same orders' destruction (either via cancellation or trade). This measure is used in the examination of order lifetime distributions for cancelled and traded orders. Because there is not very much existing results to compare the order lifetime results with they are more of a supplementary analysis to help interpret the other results.

Excess hazard ratio

Based on previous empirical studies it is expected to get findings such as clustering of events and a semi-persistent intra-day periodicity pattern of event arrivals (E.g Hall and Hautch 2004, Zhao 2010). This poses an interesting question[10]: To what extent the clustering can be explained by the (clock-time) periodicity alone? I.e. is the observed clustering caused by the concentration of events taking place at predictable intraday times? To examine this, I create a simulated arrival data for each e_1 and for each day of real data by creating a ECDF of the 20 nearest trading days data of the given stock and event, and then drawing from it the same number of observation as there is in the real data of that day. The simulated primary event times are used to create an inter-event hazard rate with real secondary event data exactly as explained before (equation 3.7) to get the *simulated empirical hazard rate function* (SHRF) for which I use the notation: $\mathrm{SHR}_{e_1,e_2}(x, c, d)$. To measure how much clustering there is in comparison to what is expected from the intraday periodicity of the events alone, SHRF can be used as a baseline to which I compare the real observed empirical hazard rate. To do exactly this I define *excess hazard ratio* (XHR) as:

$$\mathrm{XHR}_{e_1,e_2}(x, c, d) = \frac{\mathrm{EHR}_{e_1,e_2}(x, c, d)}{\mathrm{SHR}_{e_1,e_2}(x, c, d)}. \tag{3.8}$$

One clear benefit of working with XHR is that it enables the direct comparison between different calendar time partitions and companies data (something that is obviously not true for the original EHR). Additionally, XHR is also particularly convenient to interpret. E.g $\mathrm{XHR}_{e_1,e_2}(x, c, d) = 2$ means that for company c in calendar time partition d it is estimated twice as likely for e_2 to occur x ms after

[10] The question was raised before by Tuominen (2012). However, he didn't have a chance to examine it empirically.

e_1 than predicted by the intraday periodicity of the events alone given that it has not happened by that time. On the other hand, a flat function value of 1 for all values of x would indicate that after accounting for clock time periodicity there is no clustering between the events.

Response peak and mean

To determine the strength and speed of algorithmic response in the order book I use the response peak and the time of the peak as a proxy. I find it by taking the local maximum at $x \in [10..40]$[11]. The peak size is then given by

$$\text{PEAK-XHR}_{e_1,e_2}(c,d) = \max_x(\text{XHR}_{e_1,e_2}(x,c,d)), \quad x \in [10..40], \tag{3.9}$$

and the peak location:

$$\text{PEAK-x}_{e_1,e_2}(c,d) = \arg\max_x(\text{XHR}_{e_1,e_2}(x,c,d)), \quad x \in [10..40]. \tag{3.10}$$

In addition I measure the mean of XHR troughout the entire window given by:

$$\text{MEAN-XHR}_{e_1,e_2}(c,d) = \frac{\sum_{x=1}^{2000} \text{XHR}_{e_1,e_2}(x,c,d)}{2000}. \tag{3.11}$$

Produced hazard rate measures

The hazard rates are calculated for each data partition with each event type pair (e_1, e_2) for both real and simulated setups, for all of the four possible first-second event side combinations, and using window $\tau_{HR} = 2000$. The real and simulated hazard rates are then used to produce the excess hazard ratios for same side events (combining bid-bid and ask-ask) and opposite side events (combining ask-bid and bid-ask. Additionally order life time hazard rates are calculated for window $\tau_{OL} = 10000$.

[11]The constraint has been defined heuristically. An obvious alternative specification would be to simply omit the constraint, but this leads to problems as there will sometimes, although rarely, be a highest local maximum around e.g. 1000 ms, even though that is obviously not first response peak.

3.2.3 Price impact measures

I study the short-term conditional price impacts associated with the primary event types to describe the practical effects and economic significance of the event periodicity and clustering. My method is similar to that of Degryse et al. (2005), but it measures the price impact in clock-time[12].

Quote price impact

Consider the i:th event of type e and company c in d:th calendar time partition, with timestamp $t_{\text{original},e}(i,c,d)$. Then reduce the original $n_e(c,d)$ event timestamps to $u_e(c,d)$ unique timestamps $t_{\text{unique},e}(j,c,d)$ where $j \in \{1,...,u_e(c,d)\}$ indexes and which have intra-millisecond total event quantities

$$\Omega_e(j,c,d) = \sum_{i, t_{\text{unique},e}(j,c,d) = t_{\text{original},e}(i,c,d)} \omega_e(i,c,d). \tag{3.12}$$

Suppose $\Omega_e(j,c,d)$ belongs to the the q:th 6-quantile of the partition, i.e.

$$Q_e(\frac{l-q}{6},c,d) > |\Omega_e(j,c,d)| \geq Q_e(\frac{q}{6},c,d), q \in 1,...,6, \tag{3.13}$$

there is an associated event time that can be written using new indexing to include the quantile information: $t_{unique,e}(j,c,d) = t_e(p,q,c,d)$. Now a corresponding symmetrical price impact time window can be defined as:

$$t_{\text{PI}}(p,x,q,c,d) = t_e(p,q,c,d) - \tau_{\text{PI}} + x - 1, \tag{3.14}$$

where $x \in \{1,2,...,2\tau_{PI}+1\}$ indexes the window and τ_{PI} is a parameter controlling the length of the symmetrical window. Then the *clock-time quote price impact* can be measured by the windows best ask and bid prices $a(t_{PI}(p,x,q,c,d),c)$ and $b(t_{PI}(p,x,q,c,d)),c)$. In the empirical analysis the observations of the best quotes are scaled by dividing them with the mid-price at the time of the event and aggregated to get price impacts $x - \tau_{\text{PI}}$ milliseconds from the event using the equations

[12] I want to be able to compare the hazard rate results on event clustering to the price impact results, and I also think it is important to understand the millisecond level phenomena created by the HFT systems. For these reasons abandon the conventional event-time in favour of clock-time. However, it is still possible to compare the size and, with some restrictions, shape of the effect with the results of earlier studies that have used the event-time setup. The clock-time setup requires that I consider multiple events of same type that occur during the same millisecond as a single event.

$$\mathrm{QPI}_{\mathrm{A},e}(x,q,c,d) = \frac{1}{u_e(c,d)} \sum_{p=1}^{u_e(c,d)} \frac{a(t_{\mathrm{PI}}(p,x,q,c,d),c)}{m(t_e(p,q,c,d),c)}, \tag{3.15}$$

where the measure is the *quote price impact* on the *ask* side for event e and,

$$\mathrm{QPI}_{\mathrm{B},e}(x,q,c,d) = \frac{1}{u_e(c,d)} \sum_{p=1}^{u_e(c,d)} \frac{b(t_{\mathrm{PI}}(p,x,q,c,d),c)}{m(t_e(p,q,c,d),c)}, \tag{3.16}$$

where the measure is the quote price impact on the *bid* side for event e. Additionally the *mean quote prce impact* is given by:

$$\mathrm{QPI}_{\mathrm{M},e}(x,q,c,d) = \frac{\mathrm{QPI}_{\mathrm{B},e}(x,q,c,d) + \mathrm{QPI}_{\mathrm{A},e}(x,q,c,d)}{2}. \tag{3.17}$$

The mean quote impact represents the impact on the mid price, and so by definition: if $x - \tau_{\mathrm{PI}} = 0$, then $\mathrm{QPI}_{\mathrm{M},e}(x,q,c,d) = 1$.

Trade price impact

Another specification, the *trade price impact.* I include the TPI as a supplementary analysis, even though I won't be able to make direct comparisons to results in the literature since typically, the realised trade prices are not considered when calculating price impacts. The measure is calculated from realized trade prices $p_{TR}(k,c,d)$ and quantities $\omega_{TR}(k,c,d)$, of trades in the window, i.e.: $t_{TR}(k,c,d) = t_{\mathrm{PI}}(p,x,l,c,d)$. The trade prices are scaled with the mid-price at event time and aggregated using volume weighted average price for both side trades separately.

$$\mathrm{TPI}_{\{\mathrm{A,B}\},e}(x,q,c,d) = \frac{1}{\sum_{p=1}^{u_e(c,d)} \sum_k \omega_{TR}(k,c,d)} \sum_{p=1}^{u_e(c,d)} \sum_k \frac{p_{TR}(k,c,d)\omega_{TR}(k,c,d)}{m(t_e(p,q,c,d),c)} \tag{3.18}$$

, where the included trades are filtered by:

1. $t_{TR}(k,c,d) = t_{PI}(p,x,q,c,d)$, and
2. $\omega_{TR}(k,c,d) < 0$ if calculating $\mathrm{TPI}_{\mathrm{A},e}$or,
   $\omega_{TR}(k,c,d) > 0$ if calculating $\mathrm{TPI}_{\mathrm{B},e}$.

Also, the ask and bid price side impacts can be combined to get the measure of the volume weighted average price of trades *(weighted) mean trade price impact* $\mathrm{TPI}_{\mathrm{M},e}$ which can be calculated using equation 3.18 without the filtering rule 2.

For example $\text{TPI}_{\text{B},e}(x, q, c, d)$ is the event time mid price relative, volume weighted average price of trades that have been executed against the bid side[13] at $x - \tau_{PI}$ milliseconds distance of the event e. Aggregation is done over all of the events in the data for company $c \in [1..21]$, calendar time partition $d \in [1..7]$ and event e quantity quantile $q \in [1..6]$.

The price impact measures for each are created for each company with each primary event type using data from each calendar time partition, all 6 primary event quantity quantiles and 2 sides separately, using window length $\tau_{PI} = 4000$.

Instant and initial price impact

The immediate price impact as defined by Bouchaud et al. (2008) is simply the change form -1 ms to 0. Additionally I define the the *initial price impact* as the difference between the price impact measure (QPI or TPI; ask, bid or mid/average price) at time $x = -1$ and mean of the same price impact measure at times $x_{IW} - \tau_{PI} \in [101..200]$. For the initial ask side[14] is given by the equations:

$$\text{IQPI}_{\text{A},e}(q, c, d) = \text{MEAN}(\text{QPI}_{\text{A},e}(x_{IW}, q, c, d)) - \text{QPI}_{\text{A},e}(-1, q, c, d), \tag{3.19}$$

and

$$\text{ITPI}_{\text{A},e}(q, c, d) = \text{MEAN}(\text{TPI}_{\text{A},e}(x_{IW}, q, c, d) - \text{TPI}_{\text{A},e}(-1, q, c, d)). \tag{3.20}$$

Initial price impact measure represents the short term change price associated by with the event. The value lagging window boundaries of $101 - 200$ are somewhat arbitrary as they have been chosen looking at the results so that the initial reaction precedes them and the window length is long enough so that individual, noisy values (especially in the case of TPI) would not affect the result too much.

Lagging and permanent price impact

I also define the *lagging price impact* as the difference between the initial price impact and the average of the price impact measure at $x_{LW} - \tau_{PI} \in [3901..4000]$

$$\text{LQPI}_{\text{A},e}(q, c, d) = \text{MEAN}(\text{QPI}_{\text{A},e}(x_{LW}, q, c, d)) - IQPD_{\text{A},e}(q, c, d) \tag{3.21}$$

[13] meaning that the market order is a sell order and the pre existing limit order is a buy order

[14] Just ask side is given here but it is calculated bid side and mean price impacts price impacts exactly the same

$$\mathrm{LTPI}_{\mathrm{A},e}(q,c,d) = \mathrm{MEAN}(\mathrm{TPI}_{\mathrm{A},e}(x_{LW},q,c,d)) - ITPD_{\mathrm{A},e}(q,c,d) \tag{3.22}$$

Lagging price impact measure represents the more gradual drift after the initial impact, and the sum of initial and lagging impacts is a proxy for the *permanent price impact* as defined by Bouchaud et al. (2008).

3.2.4 Aggregation and partitioning

One objective of the analysis is to study how well the results generalise across the calendar time partitions, different companies and, in the case of price impact, the relative size (quantity) of the event. This section describes how different aggregated measures are formed.

Aggregation

I aggregate the different measures results based on company, calendar time, and relative event quantity. I do this by creating the measure again from the concatenated set of underlying observations.

In the case of the periodicity patterns, the data bins of different partitions are simply summed together. For the hazard rate measures I gather all of the included companies and calendar time partitions real and simulated event time differences to calculate the aggregated hazard rate measures EHR and SHR, and then get the aggregated XHR. Furthermore, I aggregate the empirical hazard rates of four different side combinations (bid-bid,bid-ask,ask-bid,ask-ask) of the primary and secondary events into two measures concatenating same side (bid-bid,ask-ask) and opposite side (bid-ask,ask-bid) event time difference observations into one measure. This is done because it seems (based on my data and previous studies) that same side measures follow the same dynamics as do the other opposite side events so further division of the measure would not yield any new information. An exception to the former is the price impact measures: I follow the same concatenated observations method to first aggregate along any other dimension and then take a simple mean of the measures company-wise.

Comparison of the measures across different data partitions

Additionally, I calculate results comparing the measures created for different data partitions. E.g in the analysis the periodicity measure is created by aggregating over all companies but splitting the data based on the calendar time partitions.

This allows the study of the pervasiveness of the periodicity pattern, as I can then compare the measures of different calendar time partitions, companies to see how much they differ or if there are clear trends.

3.3 Results

This section presents the results of the empirical analysis. The first section deals with the periodicity measure, the second, the hazard rate measures, and the third, the results of the price impact measures.

3.3.1 Periodicity

The empirical results on clock-time periodicity contain the distributions of primary event types (TR, C1, S1) timestamps compared to the level that would be expected if the events were uniformly distributed. Figure 1 illustrates the primary events periodicity patterns using all data. Overall the results show that the different event types periodicities are highly correlated regardless of the window length. Note that this does not indicate a similar absolute amount of events between the event types because the values are relative.

Subplot A contains intra-second normalized proportions within 1 millisecond wide bins ($\tau_P = 1000$,$s_P = 1$). The intra-second results are qualitatively similar to those of Hasbrouck and Saar (2013), when they study the NASDAQ (US) companies event arrival periodicity using data from 2007 and 2008. First, the arrivals are concentrated at the start of the second. Second, there are several spikes and mounds that are located slightly after certain round timestamps such as 0 ms and 500 ms. Third, the large, sudden spikes are often followed by a gradual slope and a set of smaller gradually diminishing spikes. As noted by Hasbrouck and Saar (2013) these effects are most likely caused by AT strategies that make periodic checks to the market conditions (e.g.,. at the start of each second) and as a result end up also updating their strategy and taking action (submitting or cancelling orders) in a periodic cycle. These patterns end up being amplified by the responses from other, non-periodic AT strategies. A significant difference between the two results is that while in my results the relative proportion changes in the range of about 0.9-2.4 their results range in only 0.96-1.25. However, it is possible that this difference exists due to their data set having a far larger amount of companies.

Subplot B features the intra-10-second, 10 ms bin periodicity ($\tau_P = 1000 \times 10$,$s_P = 10$). The plot reveals a regular ”heartbeat” of the order flow. Overall activity picks at the beginning of the second, and the spiky pattern repeats retaining most of the features in each 1-second segment. This result simply serves to confirm that the observed intra-second effect is something that keeps repeating from second to second as opposed to being produced, for example, as a result of large spikes and unevenness in event count during just the first second of some larger interval (e.g., one minute) followed by more flat arrival density during the rest of the interval.

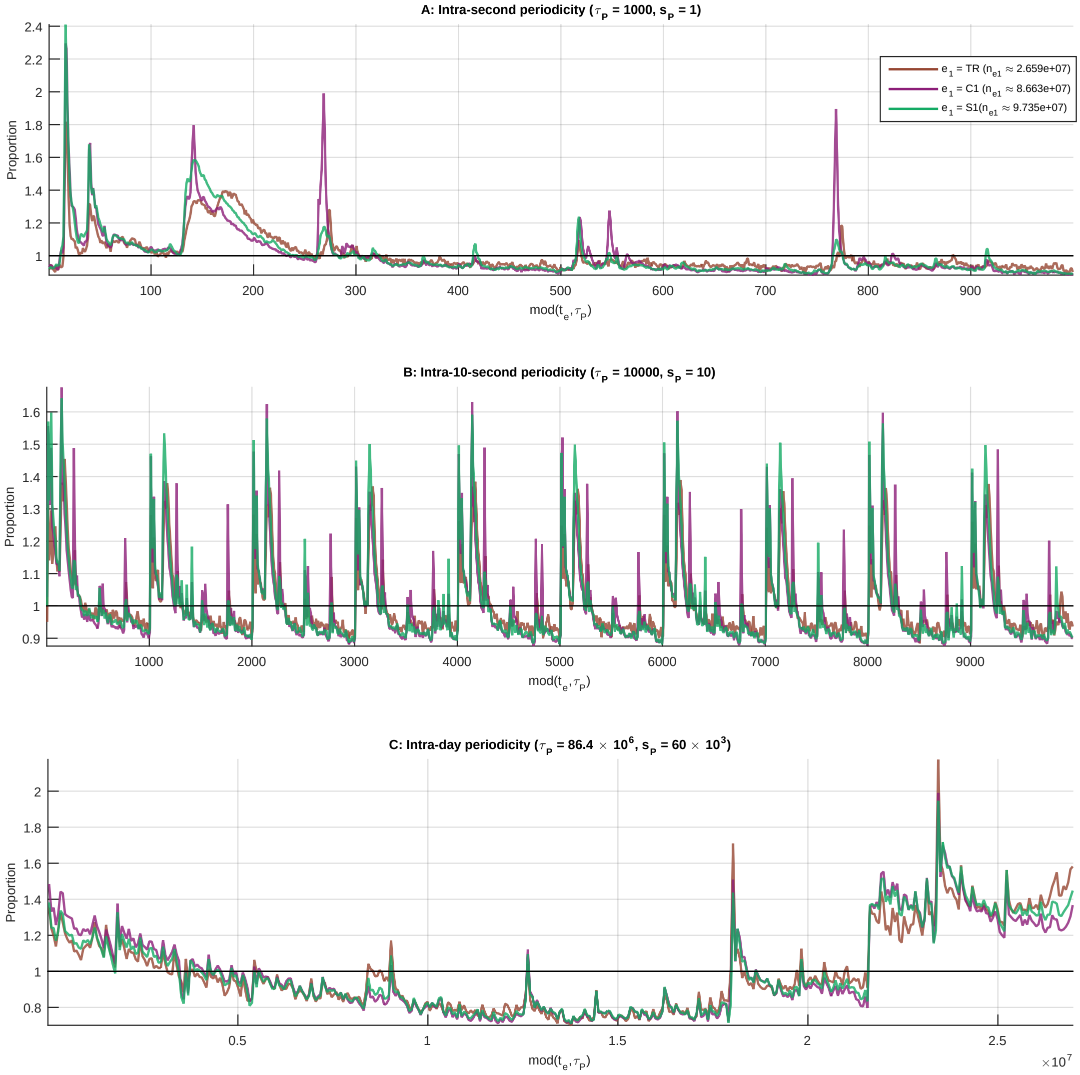

Figure 1: Aggregated event arrival periodicity patterns for trades, best level cancellations and best level submissions using all data. The y-axis is normalised by multiplying by the number of bins so that horizontal line representing the uniform distribution (no periodicity) is at 1. Plots are based on the timestamps of roughly 27 million trades (TR, orange), 87 million best level cancellations (C1, purple) and 97 million best level submissions (S1, green). The measures have been created by combining all observations form from 21 companies and 7 date partitions. The subplots represent periodicity (see equation 3.1) measures at different choices of window width, τ_P, and bin width, s_P. Subplot A contains intra-second ($\tau_P = 1000$) normalised proportions within 1 millisecond wide bins ($s_P = 1$). Subplot B contains intra-10-second ($\tau_P = 10000$) normalised proportions within 10 ms bins ($s_P = 10$). Subplot C contains intra-day ($\tau_P = 1000 \times 60 \times 60 \times 24$) normalised proportions within one-minute bins ($s_P = 60000$).

Finally, Subplot C illustrates intra-day periodicity with one minute bin width ($\tau_P = 1000 \times 60 \times 60 \times 24, s_P = 1000 \times 60$). While these results are not the main point of the analysis they, have been created to provide a comparison point for the scale of the variation in the other subplots. The scale of the intra-second, millisecond to millisecond periodicity pattern is roughly equivalent to that of intra-day, minute to minute pattern. What I mean by this is that both contain peaks and valleys that significantly deviate from the uniform distribution baseline, and scale of these deviations is equivalent. Importantly, provided that the intra-day periodicity pattern is something academics, traders and execution algorithmic designers need to take into account then the intra-second pattern is likely of equal importance.

Another conclusion that can be drawn from the intra-day periodicity pattern is that the minute resolution shows plenty of detail that will be lost if the pattern is estimated at e.g. 1/2 hour bins as Eskelinen (2015, p. 19) does.

Calendar-time-partition-wise variance in periodicty

To examine how persistent the periodicity pattern is I split the dataset between the seven 100 trading day periods and calculate the pattern for each one separately. Figure 2 presents the results. Each subplot shows the maximum median and minimum for a single primary event type and periodicity setup.

The minimum and maximum move mostly quite tightly together indicating that the observed periodicity is for the most part persistent in my sample. The intra-second results are consistent with the findings of Hasbrouck and Saar (2013). They study the periodicity pattern for at two separate times and find that the patterns closely resemble each other. However, certain specific parts don't seem to persist over time such as the three spikes in the maximum of cancellations right after 500 ms (see subplot B1). Similar to previous sections, in order to give a tangible result, I use the intra-day periodicity as a baseline to compare the intra-second periodicity with an conclude that: In the sample the intra-second periodicity persistence is roughly equivalent to the intra-day periodicity persistence, as neither have a large max-min differences.

Company-wise variance in periodicity

Figure 3 contains the maximum median and minimum for each primary event type - periodicity setup pair when data is split between the 21 companies. The minimum and maximum data for trades (subplots A1-A3) are likely noisy because there are companies for which there are not that many trades. The subplots depicting cancellations and submissions (B1-C3), on the other hand, indicate that the intra-second

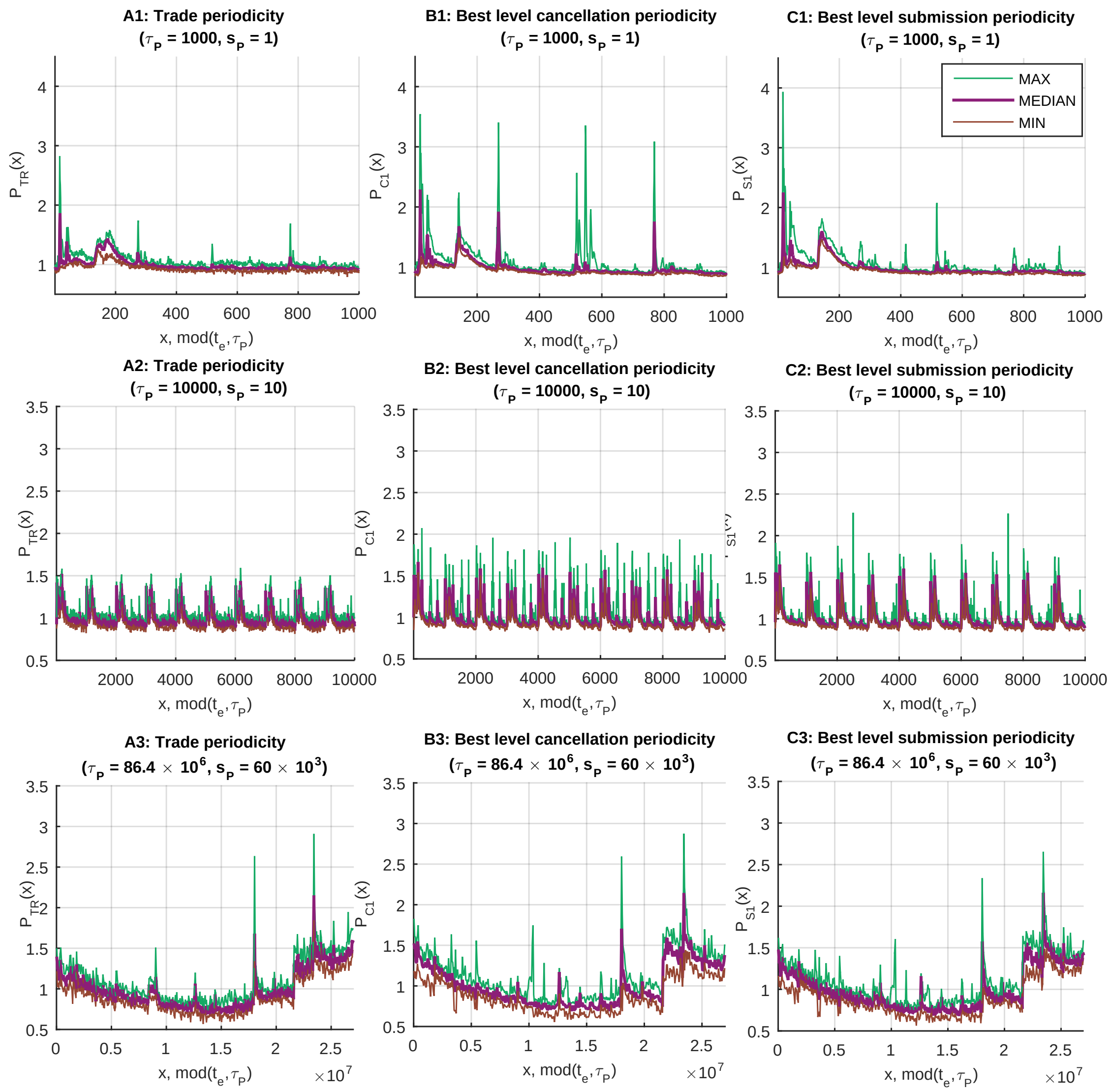

Figure 2: Calendar-time-partition-wise variance in periodicity. Minimum, maximum and median of event arrival periodicity patterns for trades, best level cancellations and best level submissions. The subplots represent distributions of 7 different calendar time partitions periodicity measures given by equation 3.1. Varying vertically there are different choices of window width, τ_P, and bin width, s_P. From top to bottom they are ($\tau_P =$ 1000,$s_P = 1$), ($\tau_P = 1000 \times 10$,$s_P = 10$) and ($\tau_P = 1000 \times 60 \times 60 \times 24$,$s_P = 1000 \times 60$). Varying horizontally there are different choices event, e. From left to right trades (TR), best level cancellations (C1), best level submissions (S1).

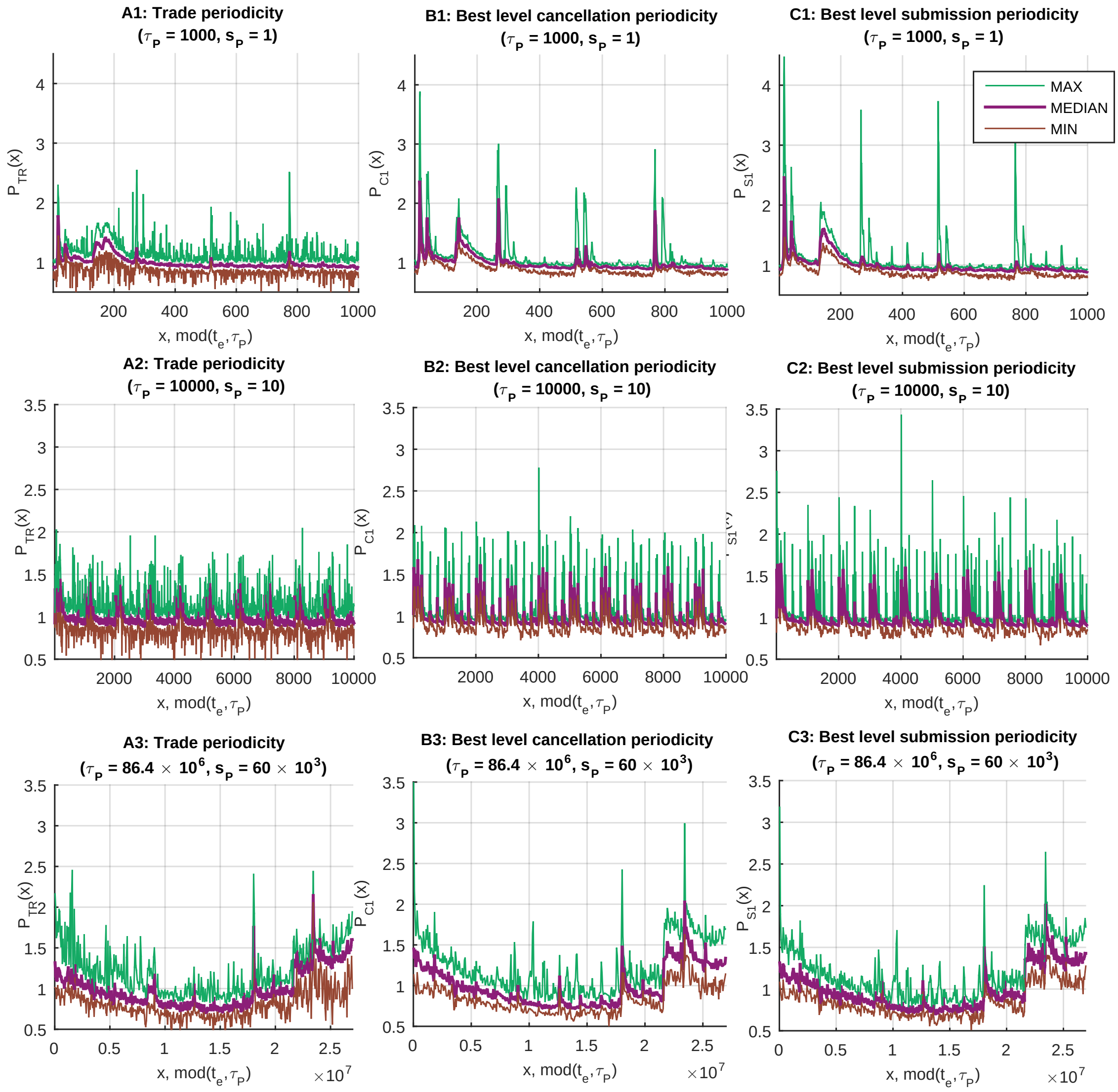

Figure 3: Company-wise variance in periodicity. Minimum, maximum and median of event arrival periodicity patterns for trades, best level cancellations and best level submissions. The subplots represent distributions of 21 companies periodicity measures given by equation 3.1. Varying vertically there are different choices of window width, τ_P, and bin width, s_P. From top to bottom they are ($\tau_P = 1000$,$s_P = 1$), ($\tau_P = 1000 \times 10$,$s_P = 10$) and ($\tau_P = 1000 \times 60 \times 60 \times 24$,$s_P = 1000 \times 60$). Varying horizontally there are different choices event e. From left to right trades (TR), best level cancellations (C1), best level submissions (S1).

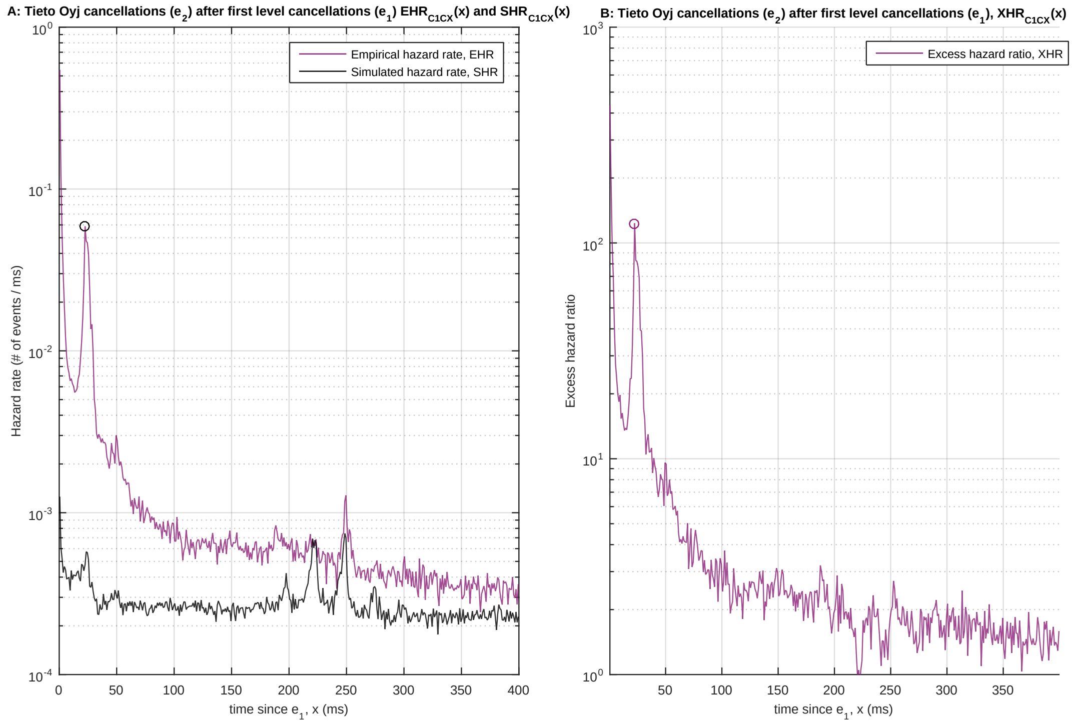

Figure 4: Example illustrating the working principle of the *Excess Hazard Ratio* (XHR) measure, which represents the level of obseved hazard rate between two event types, compared to expected hazard rate purely based on periodicity. On the left subplot there is empircal hazard rate $\text{EHR}_{C1,CX}(x, c, d)$ (equation 3.7) and simulated hazard rate $\text{SHR}_{C1,CX}(x, c, d)$ (equation 3.7) based on Tieto Oyj's entire order flows ($c = 2$, $d = [1..7]$). On the left, there is the corresponding excess hazard rate, $\text{XHR}_{C1,CX}(x, c, d)$ (equation 3.8). The measures are about any cancellation, CX, following a best level cancellation, $C1$. Additionally there is a circle marker indicating the coordinates of response peak, the location $\text{PEAK-x}_{C1,CX}(c, d)$ (equation 3.10), and peak size, $\text{PEAK-XHR}_{C1,CX}(c, d)$ (equation 3.9) in both figures (on the left one the peak is based on EHR values).

periodicity effects generalise over the companies (at least the ones studied) quite well, perhaps even better, but at least as well as the intra-day periodicity effects. This result most likely indicates that the same algorithmic trading strategies are being used in all of the sample companies' order books.

3.3.2 Hazard rate

This section covers the result relating to the hazard rate measures. Figure 4 demonstrates the idea behind the excess hazard rate measure using Tieto Oyj cancellation data from all calendar time partitions as example data. The figure shows how dividing the empirical hazard rate with the simulated hazard rate filters out some of the peaks, such as the one at 250 ms. However, other peaks such as the response peak at around approximately 25 ms is mostly unaffected. The result illustrates that the periodic pattern does not adequately explain the response peak, but rather it is in a large part created by direct cause clustering, i.e., responses to the primary event (first level cancellation in this case) and cannot be explained by clock time periodicity. As a result, there is no corresponding major peak in the simulated hazard rate.

Same side and opposite side event XHR measures created using all of the data can be seen in figures 5 and 6. The results overall are very similar shape when looking at corresponding EHR measures, which means that the contribution of clock-time periodicity is not very large as the simulated SHR measures are rather flat in comparison to EHR. However, the relative scaling of XHR is still useful and this is why I report results using it.

There is a high level of clustering in both same and opposite side measures. Depending on the event pair the same (opposite) side XHR starts off in the range 226-1324 (130-416) in the first (0-1 ms) bin, then declines rapidly to bottom at 10-42 (9-18) at 15-18 ms, turns around to form a peak of 63-329 (40-109) at around 22-24 ms and then starts a decline with multiple smaller peaks at various times that vary between event type pairs. These results indicate that there might be clustering of events (of all types) caused by exogenous factors because the XHR level is so high instantly. However, at least part of the clustering during the first couple of milliseconds is due to pegged orders being automatically moved around when either $a(t)$, $b(t)$ changes. Also starting earliest at 15 ms (where the incline of XHR starts), there seems to be a feedback effect, which is caused by voluntary algorithmic responses to the original event. Same side responses are overall stronger than opposite side for all event pairs. Additionally both the level of common cause clustering[15] and level of direct cause clustering[16] seem to be higher in the same side XHR.

Trades as primary events (orange) have with few exceptions the highest XHR with all secondary events. Also perhaps somewhat surprisingly it seems that trades on one side are associated with new trades on that side with certain event intervals such as 500, 1000 and 1500 ms. This behaviour is odd because it seems coordinated by the market order submitter and could expose the submitter to increased market impact. This is also a very strong effect. E.g in the case of 1500 ms spike, given there is still ten-fold increase of XHR even though the effect is mainly present in the first data partition and presumably not all the time even there. On the other hand, the market order sign autocorrelation, in general, is expected based on the empirical literature (Hall and Hautsch, 2004; Degryse et al., 2005; Toth et al., 2015).

[15] Proxied by level of XHR when $x < 15$ ms if we assume that very few traders can respond in that time, and the clustering before that is mostly common cause clustering or automated clustering e.g., from pegged orders.

[16] Proxied by change of XHR in 15-24ms

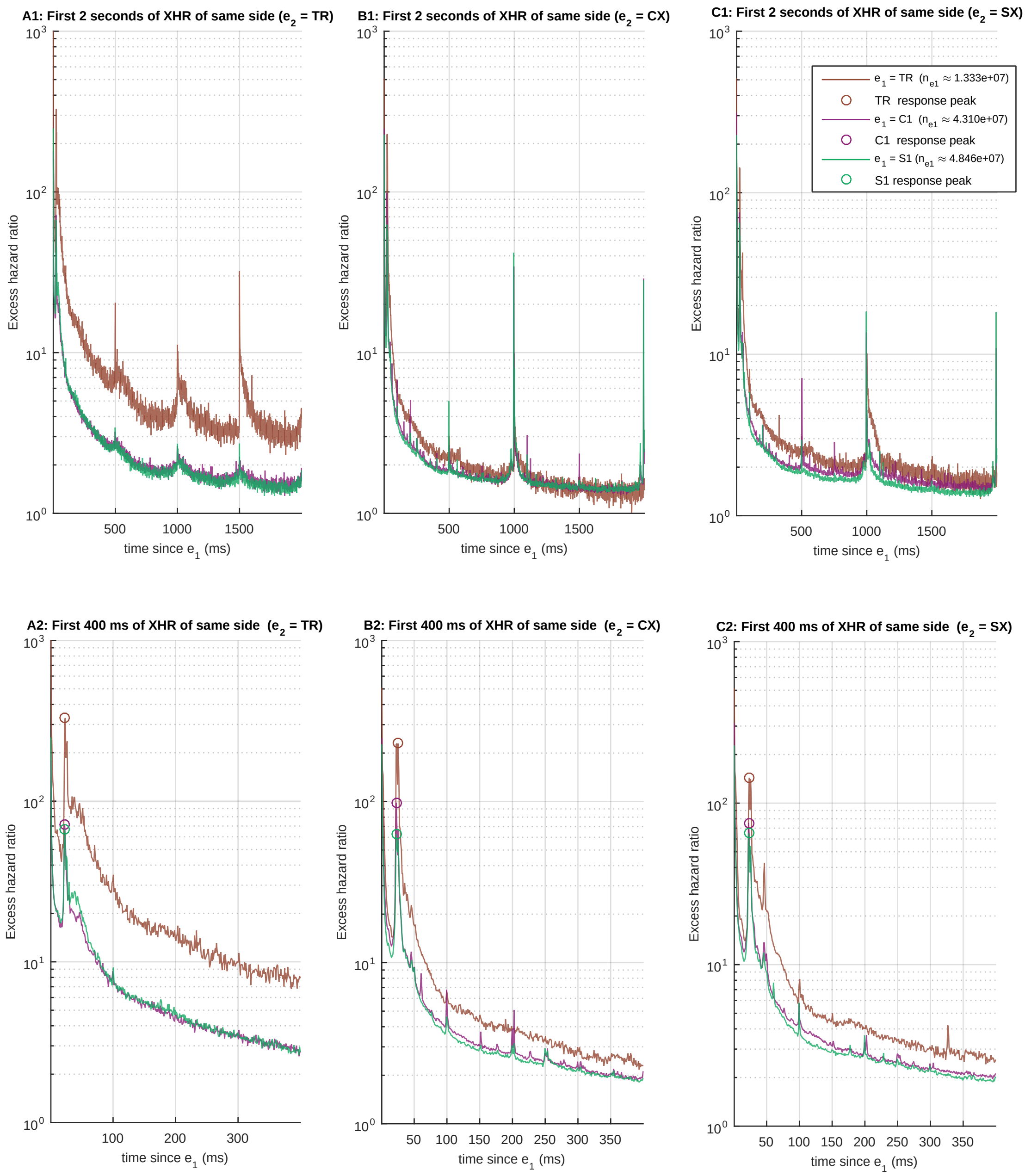

Figure 5: Excess hazard ratio between *same* side events e_1 and e_2 based on the combined inter-event times from all 21 stocks and 7 calendar time partitions. In all of the subplots, the x-axis is the inter-event time and y-axis is the corresponding excess hazard ratio value. Each subplot contains the excess hazard ratio, XHR_{e_1,e_2} (equation 3.8) of a given secondary event e_2 after primary event type e_1 is trade (TR, orange), cancellation (C1, purple) or submission (S1, green), where e_1 and e_2 are same side events. In subplots A the secondary event type e_2 is the trade (TR), in subplots B it is any cancellation (CX) and in subplots C it is any submission (SX). Upper subplots A1, B1 and C1, show the entire window up to 2000 ms after the primary event. Lower subplots A2, B2 and C2, show just the first 400 ms to illustrate the response peak (equations 3.9-3.10), which has been marked for each plot with a circle of the corresponding colour.

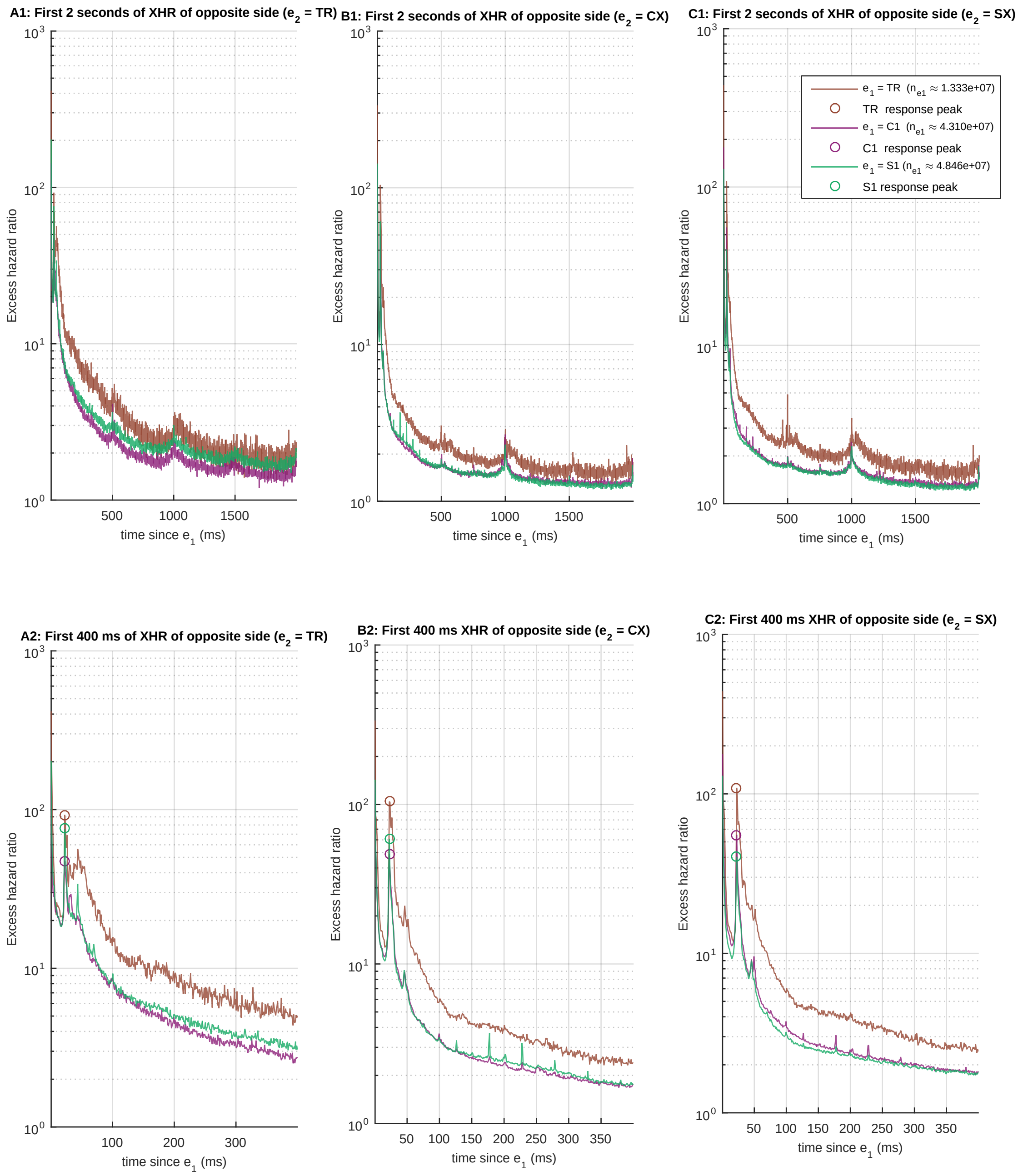

Figure 6: Excess hazard ratio between *opposite* side events e_1 and e_2 based on the combined inter-event times from all 21 stocks and 7 calendar time partitions. In all of the subplots, the x-axis is the inter-event time and y-axis is the corresponding excess hazard ratio value. Each subplot contains the excess hazard ratio, XHR_{e_1,e_2} (equation 3.8) of a given secondary event e_2 after primary event type e_1 is trade (TR, orange), cancellation (C1, purple) or submission (S1, green), where e_1 and e_2 are opposite side events. In subplots A the secondary event type e_2 is trade (TR), in subplots B it is any cancellation (CX) and in subplots C it is any submission (SX). Upper subplots A1, B1 and C1 show the entire window up to 2000 ms after the primary event. Lower subplots A2, B2 and C2, show just the first 400 ms in order to illustrate the response peak (equations 3.9-3.10), which has been marked for each plot with a circle of corresponding colour.

Furthermore, results of Toth et al. (2015) give reason to suspect (assuming that their observations apply in HSE) that is much more likely these effects are due to order splitting than herding. One explanation is that there has been agency algorithm that is executing large transactions at TWAP with almost exactly 1500 ms interval between the transactions. However, there is a possible alternative explanation, which maybe applies to at least one of the observed spikes in same side trade hazard rate. It might be that new limit orders keep getting supplied to a given price level at a particular time interval, and then those orders get immediately executed, in which case the observed pattern arises indirectly from the limit order submitters decisions. Evidence, in the case of 1000 ms spike, can be seen in subplot C1, where trades seem to trigger a resupply of the same side starting at exactly 1000 ms.

First level cancellations (purple) and submissions (green) as primary events cause rather similar responses. Both are associated with increased activity at even intervals such as 100, 200, 500, 1000, 1500, and 2000 ms. The likely explanation is algorithmic traders regularly re-evaluate the situation and update their orders, but the activity cycle can be conditional to the other events rather than clock-time periodic. Cancellations and submission also seem also to have a weak mirroring relationship in that the activity spikes tend to be higher for cancellations after submissions and vice versa on both same side and opposite side.

Response peak time and size

Using all data, the response peak can be determined to occur at the bin centred at 22,5 ms for all pairs in the opposite side XHR. In the same side XHR the peak is at 22.5 ms for event pairs (C1,TR), (S1,TR), (C1,CX), (S1,CX), (C1,SX) and (S1,SX); 23.5 ms for pairs (TR,TR), (TR,SX) and 25.5 ms for (TR,CX). This might be seen as an indication that for some traders it might take more time to process the information content of trades (See e.g.,. figure 5, subplot B2). The combinations of the pairs and same side peak sizes for are the following: (TR,TR), 329.30; (TR,CX), 72.10; (TR,SX), 66.85; (C1,TR), 229.30; (C1,CX), 98.38 ; (C1;SX), 62.62; (S1,TR), 143.50; (S1,CX) 75.41 ; (S1;SX) 65.10. The response peak seems to be at it highest with trades as the second event, and also trade as the first event. It is worth noting that this does not mean that there are more trades than other events in absolute terms but just that there are more trades compared to the typical amount of trades at the time of day when the primary events have taken place.

Calendar-time-partition-wise differences in clustering

Figure 7 contains the results of response peak time, value and mean XHR when data is split between the 7 calendar time partitions (see equations 3.9-3.11). Subplots A1-C1 (upper row) depict the response peak times of all of the event pairs. There is a clear downward trend with the times starting form the bins centred at 26.5-38.5 ms in the first calendar time partition and ending up at 23.5 ms. Depending on the event type the per annum change is between -4.8% and -19.0%[17]. This result indicates that the average response time to any event is decreasing which is possibly due to technological development among the algorithmic traders. The data suggests that the development is slowing down throughout the period, which could be because the majority of the improvement is achieved in processing time and there is a constant network latency due to the geographical distance to exchange that cannot be improved (without co-location). The OMX Helsinki stock exchanges representatives suggested that the fact that the data seems to be limited above 22 ms, might be because the network traffic behind these peaks is from London since the round trip network latency to Helsinki is exactly that 22 ms. That would mean that the processing times are, depending on the data partition, typically just a few milliseconds long which would also be consistent with the results of Hasbrouck and Saar (2013) (that in NASDAQ 2007-2008 the fastest processing times are 2-3 ms).

In subplots, A2-C2 show the evolution of the peak size and subplots A3-C3 show the 2000 ms mean through the calendar time partitions. Overall the values seem to be highly correlated. In several event pairs, there is a distinct decline from partition 1 to 4 followed by an incline from 4 to 7. However, the opposite side pairs tend to rise steadily through all of the data. Trades as secondary events generate the highest peaks and means consistently. Also, peaks and means on the same side are on average higher than on the corresponding opposite side, but not in every case.

Company-wise differences in clustering

Figure 8 illustrates response peak time, value and mean XHR when data is split between the 21 companies(see equations 3.9-3.11). The most common response time seems to be 23.5. Almost all of the response times are below 27.5 ms, but trades and especially on opposite side seem to make an exception.

I also examine the individual companies XHR and EHR patterns and find that 20 of

[17] or total change between about -11% and -39% over 2.35 years, which is the difference between the centre of the last and first calendar time partitions

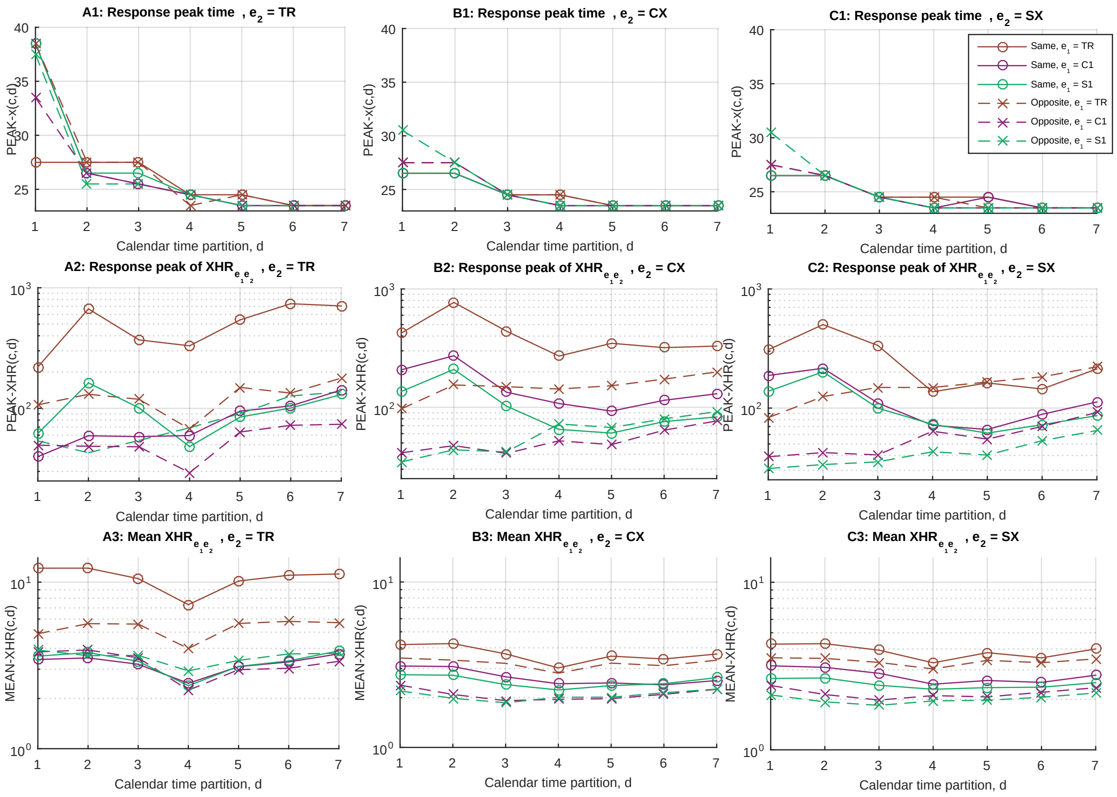

Figure 7: Calendar-time-partition-wise event response dynamics. Upper row depcits response peak location, PEAK-$x_{e_1,e_2}(c,d)$ (equation 3.10). Middle row depicts corresponding peak values, PEAK-$\text{XHR}_{e_1,e_2}(c,d)$ (equation 3.9). And bottom row contains mean XHR, MEAN-$\text{XHR}_{e_1,e_2}(c,d)$ (equation 3.11). Each measure is calculated beween each event pair e_1 and e_2 and eact subset when data is split between the 7 different calendar time partitions. In each subplot the first event, e_1 is either trade (TR, orange), best level cancellation (C1, purple) or best level submission (S1, green). Full lines and circles indicate same side events, broken lines and x-markers opposite side events. In subplots A the secondary event type e_2 is trade (TR), in subplots B it is any cancellation (CX) and in subplots C it is any submission (SX).

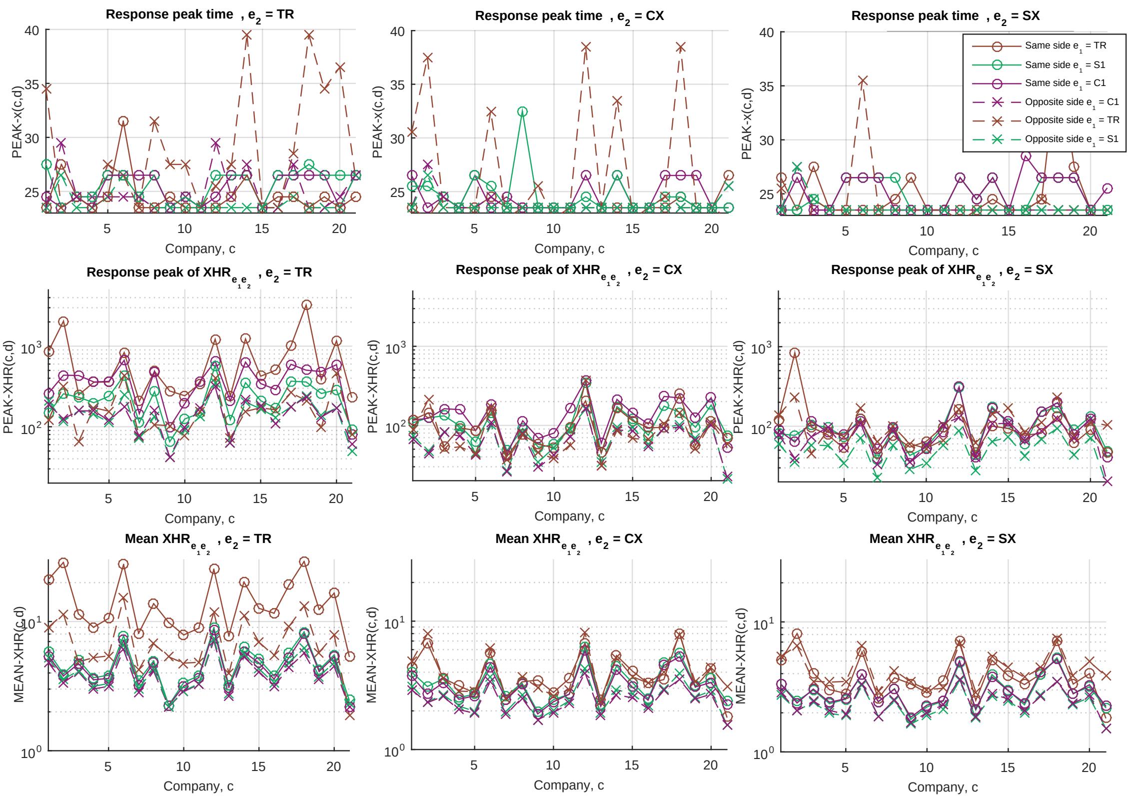

Figure 8: Company-wise event response dynamics. Upper row depcits response peak location, PEAK-$x_{e_1,e_2}(c,d)$ (equation 3.10). Middle row depicts corresponding peak values, PEAK-$\mathrm{XHR}_{e_1,e_2}(c,d)$ (equation 3.9). And bottom row contains mean XHR, MEAN-$\mathrm{XHR}_{e_1,e_2}(c,d)$ (equation 3.11). Each measure is calculated beween each event pair e_1 and e_2 and eact subset when data is split between the 21 different companies. In each subplot the first event, e_1 is either trade (TR, orange), best level cancellation (C1, purple) or best level submission (S1, green). Full lines and circles indicate same side events, broken lines and x-markers opposite side events. In subplots A the secondary event type e_2 is trade (TR), in subplots B it is any cancellation (CX) and in subplots C it is any submission (SX).

them have roughly the same shape as the aggregated figure but one (index 4, Sampo Oyj) of the companies stands out from the rest in that it has a particularly long period (up to about 130 ms) where XHR remains much higher compared to the rest. E.g, the bin at 120 ms, has the XHR value 25.23, while the median at that point is 3.78. This phenomenon might be caused by a specific set of (algorithmic) traders that trade this company especially using a unique strategy which makes these longer response times more common. Also, their 23.5 response peak is higher than usual. Thus, the result is not that the response peak moves to a different place rather there is an extended period of high XHR values after it.

Looking at the two lower rows of subplots (A2-C2, A3-C3), there is again a strong correlation in response peak and mean between all of the event pairs, which indicates that there is a company-specific factor that drives the strength of the reaction in all event pairs. Since according to Tuominen (2012) the order books are of large companies liquid stocks in Helsinki Stock Exchange during 2010 are dominated by a few HFTs (see 2.4) the company (and calendar time) -wise differences could well be explained by differences in HFT activity. To examine whether this is the case I also directly compare the reported level of HFT proportion of order submissions by Tuominen (2012) to results of response peak size for the few overlapping companies: Stora Enso (82% HFT originated orders), UPM Kymmenen (65%) and Fortum (43%), with respective indices 9,10 and 11. Treating each calendar time partitions company specific data as independent observation and assuming that the HFT activity levels do not change I find that negative correlation between log-percentage of HFT originated orders and to be between -0.45 and -0.87 in all event pairs. The fairly strong negative correlation suggests that the HFTs are not contributing to the peak but the exact opposite, which is quite puzzling, but because of the small sample size and rather an unrealistic assumption that HFT proportions remain representative throughout the 2,5-year sample this result should be checked with a larger data set.

Order lifetime hazard rate

Figure 9 gives the lifetime hazard rate for orders that are cancelled (purple) and orders that lead to trade (orange) separately to up to 10000 ms away form the original event. The empirical hazard rate is strongly decreasing and the majority of orders are either cancelled or executed in the first few seconds. The cancellations are clustered at even timestamps such as 100, 200, 250, 300, 350 ms and all the even seconds. This might be an indication of trades' tendency to periodically evaluate the viability of an order and cancel if necessary. On the other hand many of the cancellations might be predetermined to occur after a given time has elapsed if the order has not been executed using good-until-time orders. In the cancelled orders empirical hazard rate there seems to be again a "response peak" at 23.5 ms. It might be that the market state changes while submitting the order in a way that the submitter needs to cancel the order.

The lifetimes of traded order follow a similar pattern to the cancelled ones, but there are fewer peaks. There is again a noticeable peak at 23.5 ms, and the response seems to continue to up to 45 ms, after which a rapid decline starts. After that, there are pronounced peaks at 500, 1000, 1500, 2000 5000 ms and on other even seconds to a lesser extent. Here it seems somewhat odd that these peaks (other than the response peak) exist, after all, it should not matter to the market order maker how long the order has existed. However together with the result that trades seem to follow each with the same interval (see figure 5) these results can be explained by that there

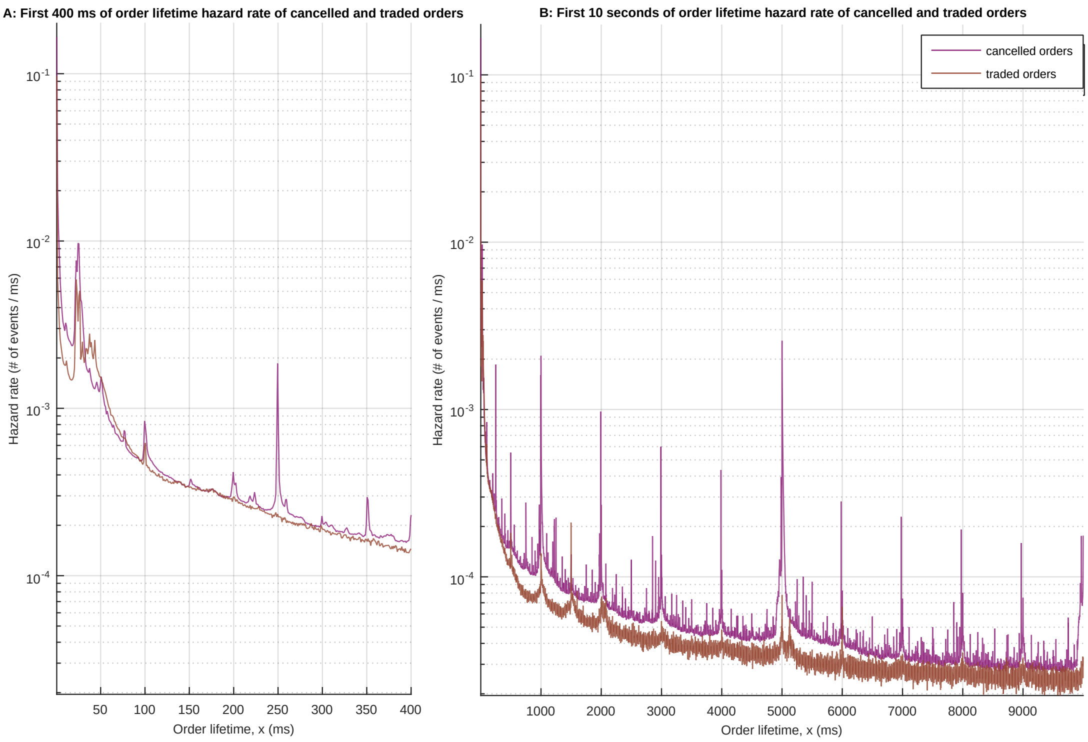

Figure 9: Aggregated order lifetime hazard rate of orders that were traded (orange) and cancelled (purple). Subplot A contains the first 400 ms and subplot B contains the entire 10000 ms window. The x-axis is time between the events that create and destroy an order. The y-axis is the corresponding value of either cancelled orders empirical hazard rate, $EHR_{LTC}(x, c, d)$ (purple) traded orders empirical hazard rate, $EHR_{LTT}(x, c, d)$ (orange) values based on all of the (explicitly) cancelled and traded orders. The measures are over all 21 companies and all 7 calendar time partitions.

are (agency) algorithms that are executing a series of trades with specific waiting time, and that the liquidity suppliers are nearly instantly replenishing the order book with new orders that, after the waiting time, get immediately traded. That would explain why the peaks in order lifetime ramp up gradually compared to the instant jumps observed in some of the hazard rate patterns. This adds to evidence that my results are in agreement with Toth et al. (2015) and the autocorrelation of trade sign, or that same side trades have overall higher hazard rates after trades compared to opposite side, is because of trade splitting and not herding.

3.3.3 Price impact

The following section covers the result of the price impact measures described in section 3.2.3. The first part will discuss the impact associated with trades; the second part will cover the results of price impact of cancellations, and the third part addresses the price impact measure results associated with best level submissions.

Trades

Figure 10 shows the results on price impact associated with bid and ask side trades using all data. Subplots A and B depict QPI around the execution of market sell and buy orders respectively while lower subplots C and D are about TPI of market sell and buy orders. One clear result regarding the plots overall is that the bid and ask side events responses mirror each other nearly perfectly, which is why I will cover only the left most plots (market sell order), and the results can be taken as applicable to the other side trades when mirrored along the y-axis.

As expected, there is an instant quote price impact (0.02% in mid price) right during the same millisecond when the trade occurs. After that, during the next 20 ms or less, there is a fast reaction where the quotes seem to drift towards each other. Somewhat surprisingly the best ask level seems to, almost instantly (starting at 1 ms) react by drifting downwards. In total, between -100 and 100 ms the bid quote moves down about 0.035%, and the ask quote moves 0.015% in the same interval. One explanation to this could be that while there is no real algorithmic response from the order submitters, there are orders pegged to the mid or bid quote, and they are automatically moved. Finally, after 20 ms mark, the spread has grown by about 25% and there is a subtle slope on both sides that makes the mid price drift down (the direction of the original response).

Examining QPI with the entire 4000 ms span also reveals that the 1000 ms 1500 ms and 2000 ms trade clustering spikes are, for certain companies and calendar

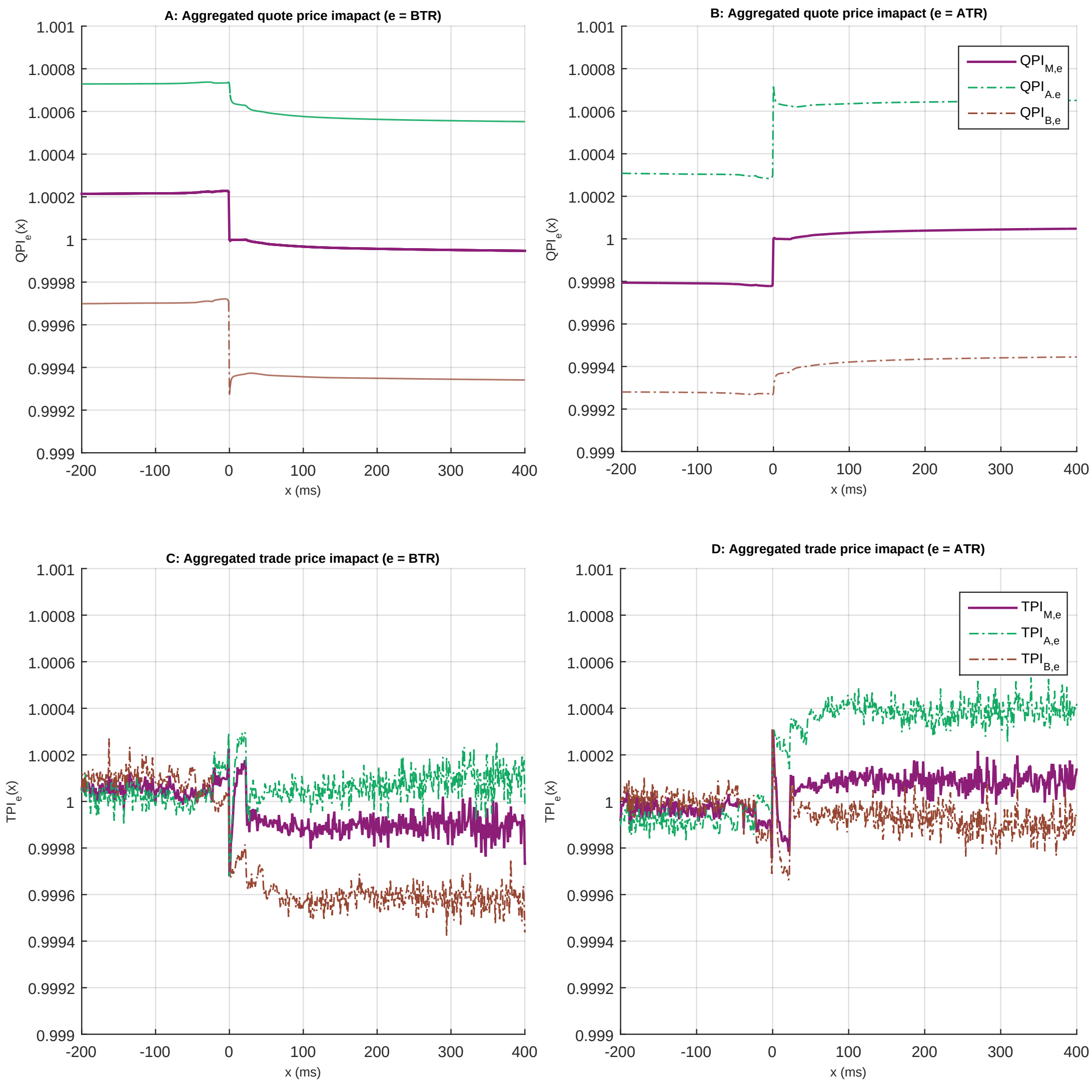

Figure 10: Market orders aggregated price impact. Quote price impact measures, $\mathrm{QPI}_{[],\mathrm{TR}}(x,q,c,d)$ (equations 3.15-3.17) are illustrated in subplots A and B and trade price impact measures, $\mathrm{TPI}_{[],\mathrm{TR}}(x,q,c,d)$ (equations 3.18) in subplots C and D. The subplots are based on aggregated data of 6 event quantity quantiles ($q = [1..6]$), 7 calendar time partitions ($c = [1..7]$) and 21 companies ($c = [1..21]$). The trade occurs when a market order is executed against bid (ask) side limit orders in left (right) side subplots.

times, visible as small but observable jumps where the price moves even more to the original 0 ms impacts direction. Especially in the first calendar time partition, there are strong 1500 ms effects and also secondary 3000 ms hump. This further confirms the finding that there are chains of several trades taking place with a minimum of 1500 ms in between.

The TPI results reveal something more unexpected. When a trade is executed against a bid side limit order, there are clear reversals at the 23 ms mark in both the ask and bid side TPI. Furthermore, on the bid side there is there are repeated reversals at 46, 69 and 93 ms. What happens is that in between these reversals the TPI moves to the opposite direction of the initial movement and at the reversals (and also slightly on both sides immediately next to them) are new quick jumps to the direction of the original movement. Also during the first about 100 ms the bid side continues to drift down for about a third of the instant 0 ms impact bringing the total change in the measure between -100 and 100 ms to about 0.05% on the bid side, but the ask side changes only about 0.005%. This phenomenon is somewhat tricky to interpret, but we do know the aggregated results for the QPI do not exhibit these jitters so the cause must be the changing ratio between executed volume in the cases where the bid quote is lower versus where it is higher. A possible explanation is that in cases where the quotes move a lot in response to the initial trade event there are surges in trading activity against the recently moved quote at these 23 ms intervals for a while until at the ratio settles to a level. If the 23 ms regular jitters are assumed to be caused by varying proportion of HFT activity then these results fit together with the finding of Brogaard et al. (2014) that HFTs tend to trade in the direction of the permanent price change.

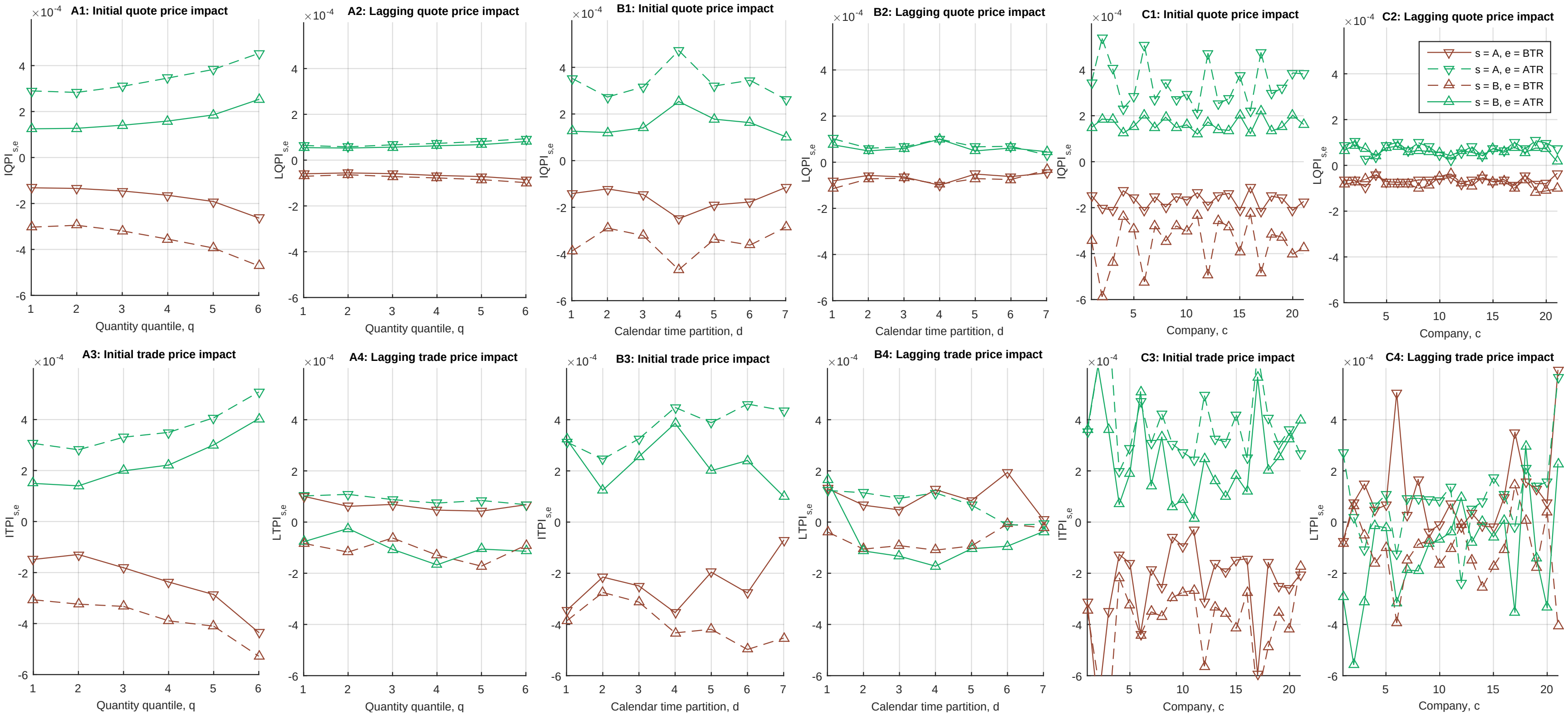

Figure 11: Event-quantity-wise ,calendar-time-wise and company-wise price impact dynamics of market orders (trades). Initial quote price impact, $\mathrm{IQPI}_{[],TR}(q,c,d)$, lagging quote price impact, $\mathrm{LQPI}_{[],TR}(q,c,d)$, intial trade price impact, $\mathrm{ITPI}_{[],TR}(q,c,d)$, and lagging trade price impact, $\mathrm{LTPI}_{[],TR}(q,c,d)$, price impacts (given by equations 3.19-3.22) of *trades* for data partitioned among different dimensions. Each subplot contains the price impact measure on each side of the book in response to an event on each side of the book. Subplots A show results for when data is split between the 6 primary event quantity quantiles, q. Similarly, data is split between the 7 calendar time partitions, c, in subplots B and between the 21 different companies stocks, c, in subplots C. Subplots [A-C]1 feature initial quote price impacts and subplots [A-C]2 contain lagging quote price impacts, while subplots [A-C]3 contain initial trade price impacts. Subplots [A-C]3 feature initial trade price impacts. Finally, subplots [A-C]4 depict lagging trade price impacts.

The figure 11 depicts the initial and lagging price impacts of for different data partitions. Green line denotes measures around an ask side event; orange line denotes measures around bid side event, downward pointing triangles denote measures on the ask side and upward pointing triangles denote measures on the bid side. Subplots A show results for when data is split primary event quantity quantile. Similarly, data is split between the calendar time partitions in subplots B and between different companies stocks in subplots C. Subplots [A-C]1 feature initial quote price impacts and subplots [A-C]2 contain lagging quote price impacts, while subplots [A-C]3 contain initial trade price impacts. Subplots [A-C]3 feature initial trade price impacts. Finally, subplots [A-C]4 depict lagging trade price.

Excluding the lagging trade price impacts in [A-C]4 there are a few general characteristics that the each subplot shares 1) Impact is always the same direction which is negative for trades executed against a bid side limit order (orange) and positive for trades executed against ask side limit orders (green). 2) Initial price impact widens the spread, i.e. the side where the trade happens experiences a larger impact than the other side, even though they both move to the same direction. 3) The bid (orange) and ask (green) side events mirror each other nearly perfectly along the x-axis, which is consistent with Zhou (2012).

The subplots A[1-4] show the results for when data is split between different primary event quantity 6-quantiles of the corresponding calendar time partition company combination. While this is not such an interesting result by itself, it is something easy to understand, and it can be used as a point of comparison when assessing the other results in this figure. The larger the primary event size, the larger the impact.

Results of calendar time wise split are illustrated in subplots B[1-4]. The largest impacts are in the fourth partition. With the current analysis setup, it is possible only to speculate, but it seems that the results mirror the figure 7, which has a slump in the fourth partition, meaning that its possible that the stronger clustering is in fact associated with the dampening of the price impact. Also if we assume that the clustering peak size is a decent proxy for HFT activity. Subplots C[1-4] depict the results of company specific analysis it shows that while the level of individual companies initial price impact varies the correlation between of the four QPI, and four TPI series inter-company values is quite high (depending on the series absolute values between 0.99-0.60), with the highest correlations recorder between the mirroring QPI measures (like ask side after ask event versus bid side after bid event) and the lowest between QPI and TPI measures. This means there must be an underlying company specific property that determines the initial price impact. The corresponding results for lagging price impact are similar. However the values are much lower in the case of QPI, and there is no clear structure with TPI related results.

Best level cancellations & submissions

Here all of the price impact related results for both best level cancellations and submissions will be covered. The reason for handling them together is that they are in a certain sense opposite actions to one another.

Figures 12,13 represent the results of QPI (upper row) and TPI (lower row) associated with best level cancellations/submissions at the bid (left column) and ask (right column) side respectively.

As seen the in the QPI subplots (A & B) on aggregate level cancellations, similarly to trades, associate with a price change with same sign on both sides of the book. Again similar to trades, the direction is also to move the side where liquidity was decreased away from the mid price. Submissions exhibit simlar behaviour except the direction of the price change is reversed. Regarding absolute size of the initial QPI effect both are considerably smaller compared to the trades' -0.023% with C1 at -0.010%, and S1 0.002%.

However, compared to trades' results for both C1 and S1 the TPI changes much more drastically and ask and bid side trades aggregate level prices diverge from each other noticeably more. Between the -100 and 100 ms the realised bid (ask) market order makers cost after event in bid side for best level cancellations, -0.085% (0.039%); and best level submissions, -0.041% (0.065%). This demonstrates that the realised trading right after the best level cancellations and submissions is on average much more expensive to the market order makers than moments before the event. However, the market order makers make these trades voluntarily. One possible explanation is that these effects are caused by behaviour of the limit order submitters to cancel and resubmit their orders a lot when there is a news item or stock exchange release being released, after which impatient informed traders flood the market, but it is difficult to explain why the same effect would not be reflected in quote prices as an increased spread.

In subplot C and D, which depict the TPI associated with bid and ask side cancellations there is some signs of the similar 23 ms effect. The change on both sides seems to peak at certain multiples of 23 around zero, but the effect is to weak to compared to noise to be definitive. In subplot D just before the ask side cancellation, there is a set of spikes between -100 ms to 0 ms that is difficult to explain, perhaps there are some extreme observations based on faulty data messing up the measure.

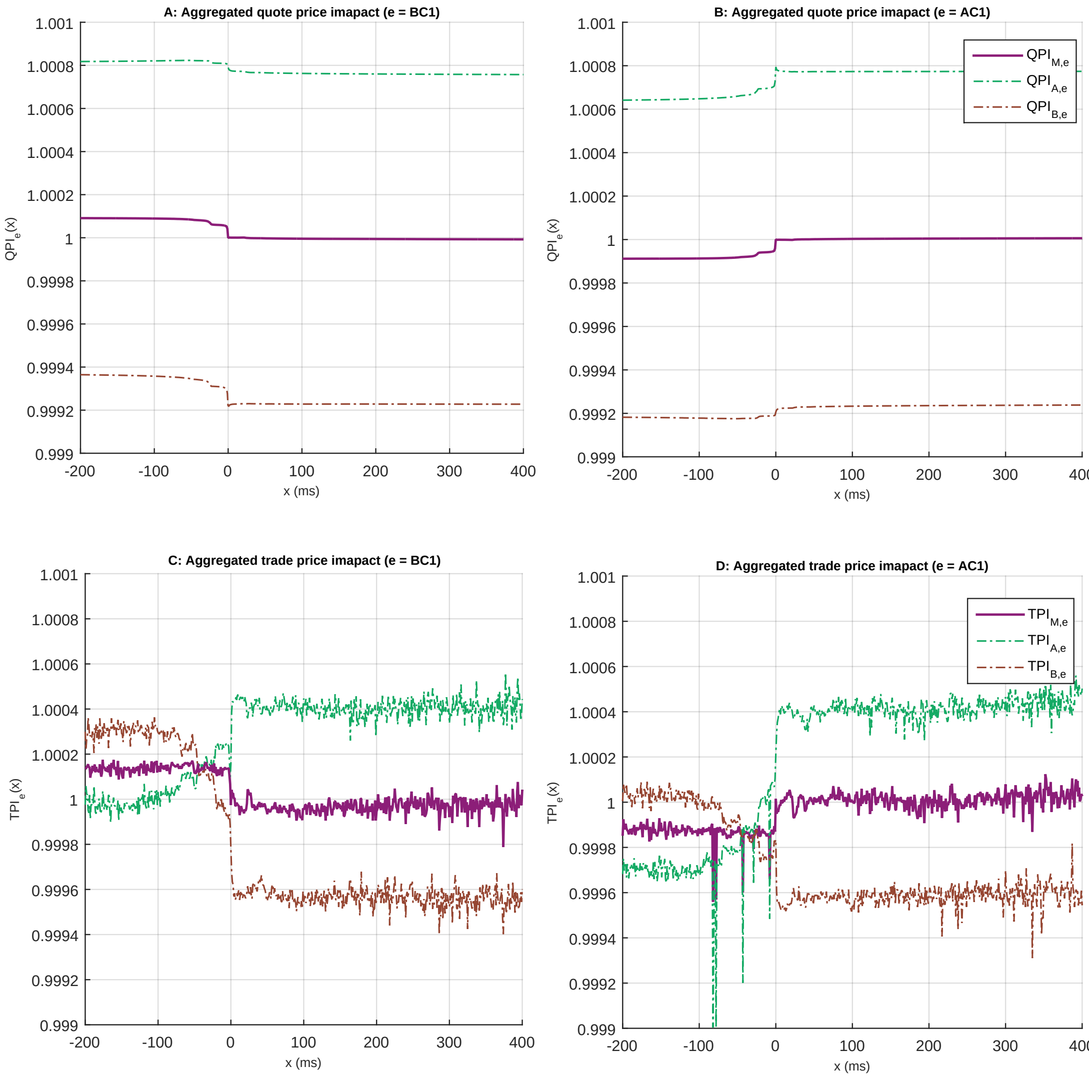

Figure 12: Order cancellations' aggregated price impact. Quote price impact measures, $\mathrm{QPI}_{[],\mathrm{C1}}(x,q,c,d)$ (equations 3.15-3.17) are illustrated in subplots A and B and trade price impact measures, $\mathrm{TPI}_{[],\mathrm{C1}}(x,q,c,d)$ (equations 3.18) in subplots C and D. The subplots are based on aggregated data of event quantity quantiles ($q = [1..6]$), 7 calendar time partitions ($c = [1..7]$) and 21 companies ($c = [1..21]$). The cancellation occurs on the bid (ask) side best level limit orders in the left (right) side subplots.

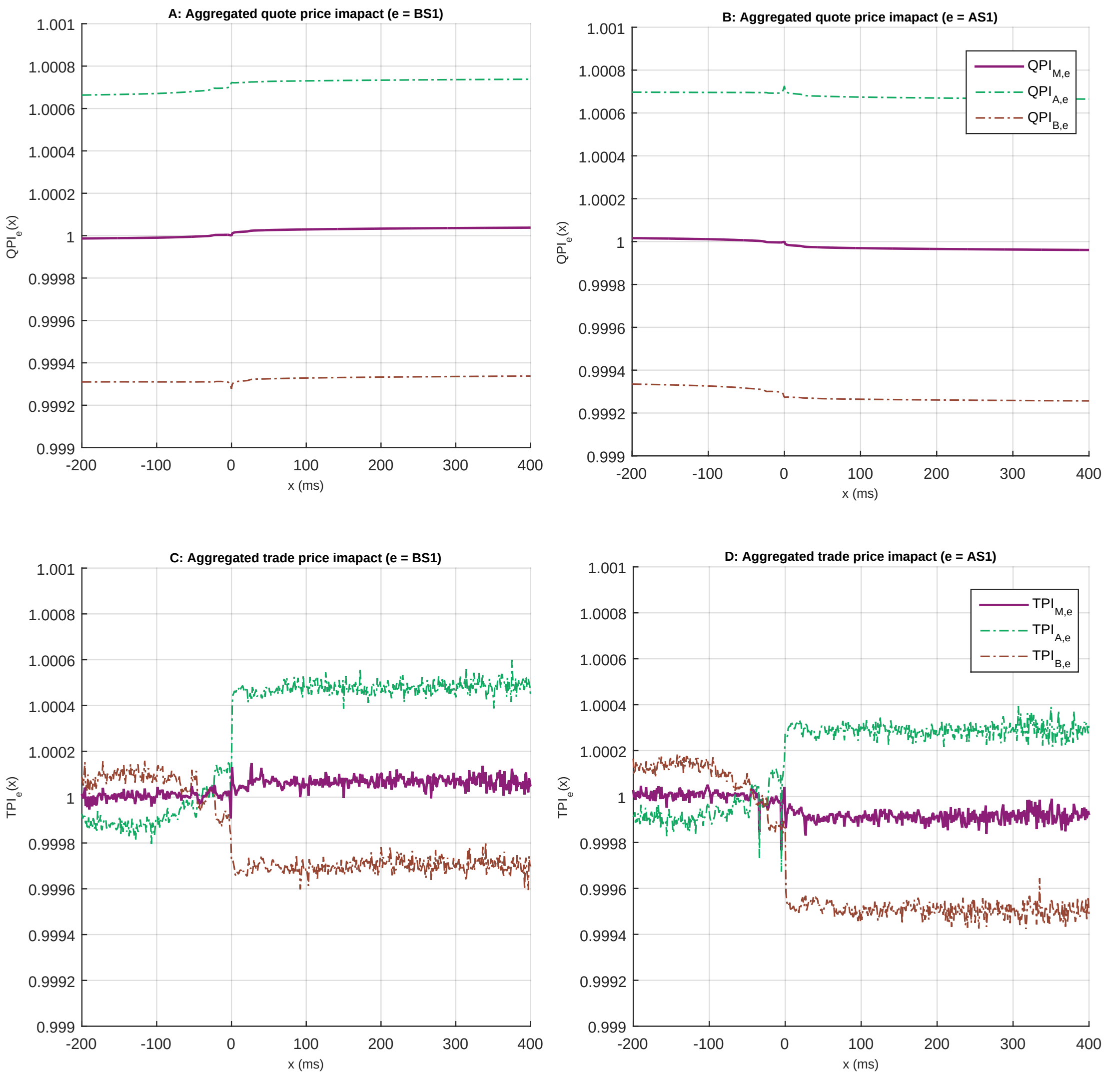

Figure 13: Order submissions' aggregated price impact. Quote price impact measures, $\mathrm{QPI}_{[],\mathrm{S1}}(x, q, c, d)$ (equations 3.15-3.17) are illustrated in subplots A and B and trade price impact measures, $\mathrm{TPI}_{[],\mathrm{S1}}(x, q, c, d)$ (equations 3.18) in subplots C and D. The subplots are based on aggregated data of event quantity quantiles ($q = [1..6]$), 7 calendar time partitions ($c = [1..7]$) and 21 companies ($c = [1..21]$). The submissions on the bid (ask) side best level can be seen in the left (right) side subplots.

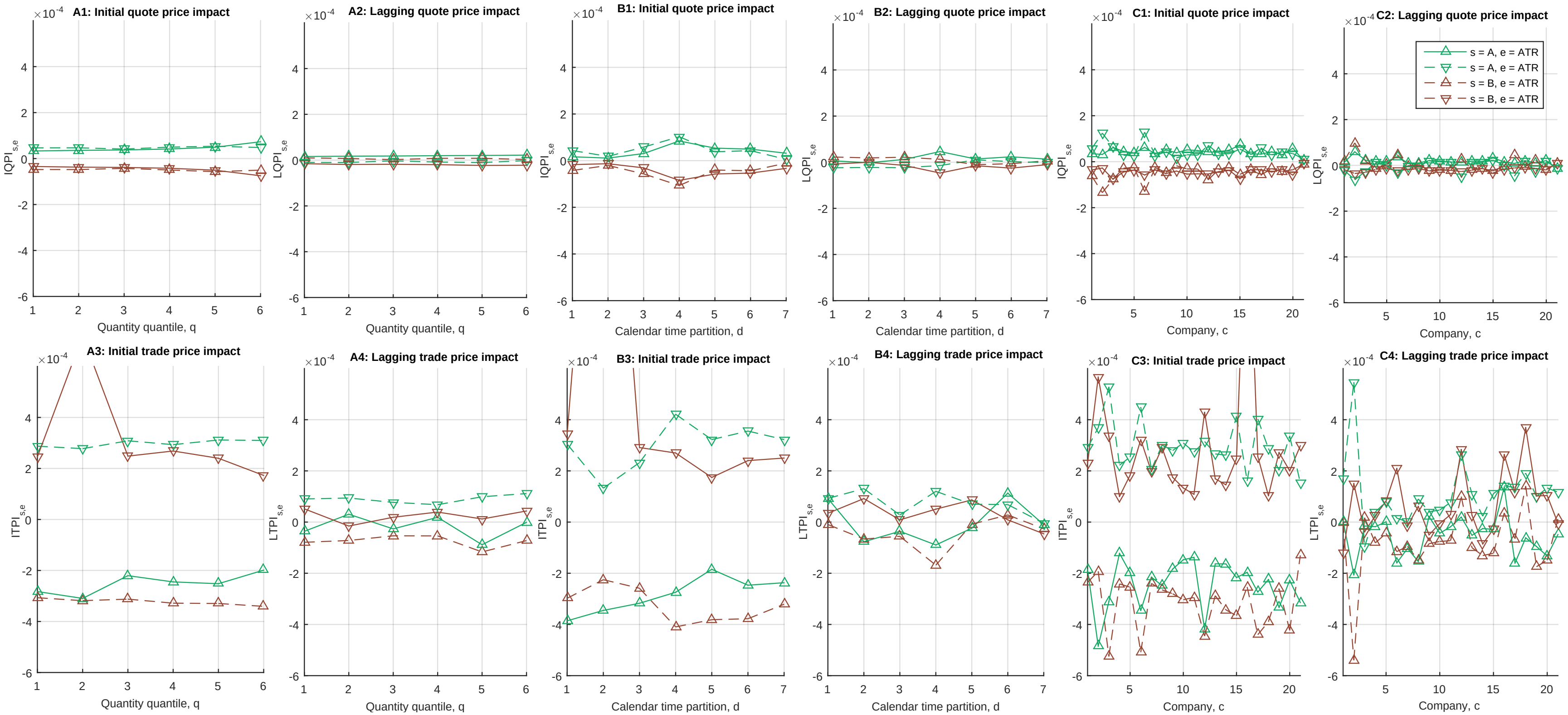

Figure 14: Event-quantity-wise ,calendar-time-wise and company-wise price impact dynamics of best level cancellations. Initial quote price impact, $\mathrm{IQPI}_{[],C1}(q,c,d)$, lagging quote price impact, $\mathrm{LQPI}_{[],C1}(q,c,d)$, intial trade price impact, $\mathrm{ITPI}_{[],C1}(q,c,d)$, and lagging trade price impact, $\mathrm{LTPI}_{[],C1}(q,c,d)$, price impacts (given by equations 3.19-3.22) of *trades* for data partitioned among different dimensions. Each subplot contains the price impact measure on each side of the book in response to an event on each side of the book. Subplots A show results for when data is split between the 6 primary event quantity quantiles, q. Similarly, data is split between the 7 calendar time partitions, c, in subplots B and between the 21 different companies stocks, c, in subplots C. Subplots [A-C]1 feature initial quote price impacts and subplots [A-C]2 contain lagging quote price impacts, while subplots [A-C]3 contain initial trade price impacts. Subplots [A-C]3 feature initial trade price impacts. Finally, subplots [A-C]4 depict lagging trade price impacts.

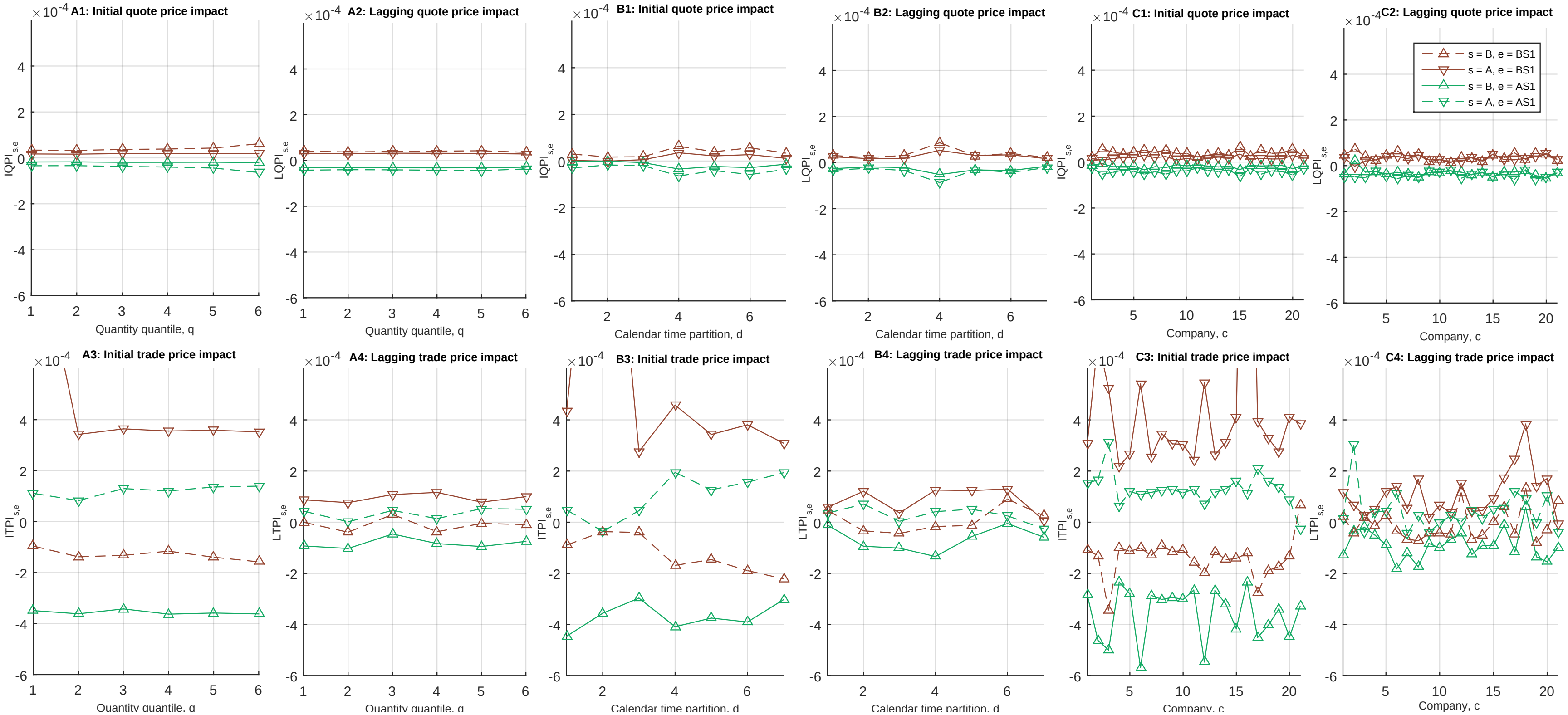

Figure 15: Event-quantity-wise ,calendar-time-wise and company-wise price impact dynamics of best level submissions. Initial quote price impact, $\mathrm{IQPI}_{[],S1}(q,c,d)$, lagging quote price impact, $\mathrm{LQPI}_{[],S1}(q,c,d)$, intial trade price impact, $\mathrm{ITPI}_{[],S1}(q,c,d)$, and lagging trade price impact, $\mathrm{LTPI}_{[],S1}(q,c,d)$, price impacts (given by equations 3.19-3.22) of *trades* for data partitioned among different dimensions. Each subplot contains the price impact measure on each side of the book in response to an event on each side of the book. Subplots A show results for when data is split between the 6 primary event quantity quantiles, q. Similarly, data is split between the 7 calendar time partitions, c, in subplots B and between the 21 different companies stocks, c, in subplots C. Subplots [A-C]1 feature initial quote price impacts and subplots [A-C]2 contain lagging quote price impacts, while subplots [A-C]3 contain initial trade price impacts. Subplots [A-C]3 feature initial trade price impacts. Finally, subplots [A-C]4 depict lagging trade price impacts.

Figures 14 (best level cancellations) and 15 (best level submissions) depict the initial and lagging price changes on both sides of the book when data is split along different dimensions. Subplots A contain results for the split based on primary event quantity 6-quantile.

Differences compared to trades are that is that cancellations are associated 1) with widening spread leading up to the event, 2) during the initial price impact spread does not notably widen or close 3) during the[18].

For both events and regardless of how the data is split the IQPI results for both sides move consistently to the same direction. Similarly, the results on ITPI are consistent (with one exception in company 21, Rautaruukki Oyj) in that the realised prices of trades diverge on different sides, i.e. ITPI of ask and bid side have opposite signs. The realised buying at market becomes more expensive and selling yields a lower price.

Quantity wise split S1 IQPI results show that the larger quantity submissions are associated with a larger decrease in the spread. Conversely, the smaller quantity cancellations seem to be associated with a growing spread, but as the relative quantity grows of larger cancellations the effect disappears, this is interesting because it is rather counter-intuitive. However the LQPI, ITPI and LTPI seem to be mostly unrelated to the event quantity, or at least there are no clear trends.

When best level cancellation and submission data is split based on date partition and companies data (subplots B and C), there are a lot of the same features what were observed with trades. QPI (and to lesser extent TPI) measures again exhibit the mirroring between the bid and ask side responses. The companies that had the largest measure values with trades also have them with best level cancellations and submissions. Also interestingly for both events, the fourth date partition seems (similar to the trades) to have the largest IQPI and LQPI effects.

[18] As these effects are rather subtle in the -200 ms to 400 ms range, I confirm this from the full -4000ms to 4000ms results

4. DISCUSSION

This part discusses the results of the empirical analysis and compares them to established empirical literature, and attempts to find explanations for the observed differences. The results are covered by addressing the research questions that were posed in section 1.1.

(1a) What are are the differences in the periodicity results between the data partitions: different calendar times and companies. For all of the studied primary events, trades, best level cancellations and best level submissions, there exists an intra-second, 1 ms bin periodicity pattern, that is 1) of comparable scale to the intra-day, 1 minute bin periodicity pattern, 2) persistent over the examined 700 day period when examined in 100 trading day parts and 3) common to studied 21 companies.

(2a) What are are the differences in the periodicity results based on my data sample compared and the established empirical results in limit order book literature. Compared to the results of Hasbrouck and Saar (2013) based on NASDAQ (US) companies in 2007-2008, my results show much higher variation in the relationship of the observed proportion and the baseline (uniformly distributed) proportion (about 0.9–2.4 vs. 0.96–1.125). However, the overall shape is similar enough to suggest there are similar mechanics behind the results. The bulk of the events are around the beginning of the second and form a right facing slope. Additionally, the measure spikes after even times such as the beginning of the second or right after the 500 ms mark.

My analysis goes beyond Hasbrouck and Saar (2013) and confirms that the patterns of all three event types closely resemble each other. Furthermore, the results on 10-second 10 ms bin also confirms that the intra-second effect is not caused by few isolated seconds (e.g., at the beginning of each minute), but rather is present all the time. Another important conclusion is that the intra-day pattern has a lot of detail when depicted in 1 minute bin resolution. Therefore it I recommend that at most 1 minute bin width is used when estimating the pattern to avoid inaccuracy caused by over-smoothing. Depending on the nature of analysis the considerable inaccuracy of such estimation can have dire consequences. Thus I recommend estimating even the intra-day periodicity using very narrow bins, at most 1 minute but preferably less. However if data is plentiful even better alternative is to use a method like the simulation utilized in creation of the XHR measures.

Although the intra-day clock-time periodicity is well known and commonly taken into account in empirical studies, the intra-second periodicity is often overlooked. With the exception of the analysis of Hasbrouck and Saar (2013), I found no references to this phenomenon or any methods used to explicitly account for this effect (of course no all analyses would even require it). In future empirical studies, it would make sense to make the distinction between clustering that is the direct result of event periodicity and clustering that is independent of periodicity, especially since the periodicity is likely to become stronger as AT and HFT activity continue to increase their share of trading activity. The next steps in research into the properties of periodicity would be to examine the evolution of the periodicity e.g. hour by hour within the trading day and expand the analysis geographically to include other Nordic markets and the US markets. Furthermore another direction would be to try to use the periodicity pattern to identify order submitters as HFT/non-HFT or informed/uninformed traders and construct trading and execution strategies based on this information.

(1b) What are are the differences in the clustering results between the data partitions: different calendar times and companies. In my analysis of the event clustering behaviour I show that in the data there is a calendar-time-wise semi-persistent and in part company-wise general hazard rate pattern, that exhibits clustering of events even if the clustering based on the observed periodicity is filtered out. There are certain common properties that all of the XHR patterns (any event pair or side) share: 1) A similar shape, a power-law-like decline, with spikes at round millisecond values. 2) A response peak, the first and usually highest[1] spike is located at soon after 22 ms. There is a clear pattern where the response time peak time approaches the bin centred at 23.5 ms as calendar time progresses decreasing at an average annual rate of -(19.0–4.8%). 3) There are additional prominent spikes at commonly located at 500 ms intervals (although some of these are found to be transient) which are in case of trades as secondary events likely signs of trade splitting rather than herding as suggested by Toth et al. (2015) 4) Trades produce by far the strongest reaction in every event type. 5) While cancellations and submissions are followed by nearly indistinguishable response except for trades which react slightly stronger following opposite submissions. 6) The events are always followed by a stronger reaction on the same side then on the opposite side of the order book[2]. 7) The response peak size and average over the 2000 ms window are quite highly corre-

[1]Highest in terms of absolute value that is.

[2]The results holds for both XHR and EHR

lated[3]. These results suggest that there is an underlying company and date partition specific factor driving the overall clustering. Since Tuominen (2012) find that in the at the start of my data set the market is dominated by handful of high-frequency traders I combine the overlapping results of HFTs proportion of submitted orders and find that the proportion of HFT activity is in fact negatively correlated (-0.45 - -0.87 depending on event pair) with the XHR response peak measure of clustering. Since this is quite unexpected and the intuitive result would have been exactly the opposite, confirmation of this result would definitely be interesting see in studies to come.

The supplementary analysis of order lifetimes for traded and cancelled orders reveals that the orders that are eventually cancelled tend to be cancelled after certain even durations after their submission, which makes sense since the submitters monitor the orders and evaluate their viability frequently using algorithms that work on even intervals. Another explanation is that traders utilise good-until-time orders. Both cancelled and traded orders lifetimes hazard rates also exhibit a spike usually at the 23-27 ms depending on the data partition.

The fact that XHR values are similarly shaped to the original EHR values, and therefore clustering cannot be explained by the clock-time periodicity alone there must be other, intermittent factors that cause the common causation clustering in the first few milliseconds where it cannot plausibly be response driven. Based on this another interesting research direction would be to try and find those exogenous drivers of the event arrival process. Public information releases like news and stock exchange releases are obvious candidates.

(2b) What are are the differences in the clustering results based on my data sample compared and the established empirical results in limit order book literature. Comparing EHR to the results of Hasbrouck and Saar (2013) the results are somewhat similar. They have a decreasing hazard rate with response peak. However, the peak is at 2-3 ms after the primary event. At a glance, the ≈ 23 ms event response seems to be at conflict with their findings. Given that they study NASDAQ (US) data from 2007-2008 it seems unlikely that in my sample, several years of technological progress after the processing times would be even at best still over 7-10 times slower. However, it was pointed out by a member of NASDAQ Nordic technical staff (in response to my inquiry into the possible origins of this phenomenon), that round trip latency to between London and Helsinki is 22 ms (11 ms per direction). Assuming that this response peak is in fact caused by London originated orders and the 22 ms is the network latency and (at least) 1 ms goes to

[3] always over 0.6 for pairs with the same event and the same side in both calendar-time-wise and company-wise split data

processing the information the results two sets of results can be reconciled. This seems a quite likely explanation as it also explains why the market response peak seems to approach the 23.5 ms bin and not breach below it. However this result is also quite puzzling as it indicates that there are London based algorithmic traders participating in the Helsinki stock exchange that either 1) choose not to co-locate for faster access or 2) need to run the decisions or a part of them by London in any case.

In any case, this result presents an interesting question about how do these kinds of ”peripheral market” conditions affect the Helsinki market, and other, similar markets. Further research might be warranted in such as e.g., 1) A confirmation of the phenomenon e.g. by looking at other Nordic exchanges and finding if they experience the same with different latency would be in order. 2) Next step would be to split the event feed into groups of likely and unlikely ”London traffic” using what ever data is available and then characterize the two groups: what is their, profitability, propensity to use limit orders and share of trading and orders. 3) It would also be interesting to try to identify the fastest responding HFTs and see if they are able to exploit the rather long 23 ms latencies by front running their counterparts.

The trade, same side trade hazard rate measure also reveal spikes at even intervals such as 1000, 1500 and 3000 ms. While the autocorrelation of trade/order size consistent with earlier studies (Degryse et al., 2005) it is surprising that there are so clear intervals, because it seems like they would expose the market order submitters plan and lead to a costly price impact even before she/he executes the entire batch. However based on the results of Toth et al. (2015) it does seem more likely that reason for the observed sing autocorrelation is, in fact, order splitting and not herding.

(1c) What are are the differences in the price impact results between the data partitions: different first event quantity quantiles, calendar times and companies.

The results on the price impact of different size trades are expected given the results of the literature. I also find a similar size relationship to apply to the best level cancellations and submissions, but it only applies only to the opposite (same) side of the book where the cancellation (submission) took place. The literature reports concave relationship between price impact and quantity, but my results cannot be tested against this result because the quantile-split bins have uneven widths. The results vary a lot between calendar times and even more between the companies. However, there is clear correlation within the variables, and it seems that the dynamics could be described by one or two parameters that quantify the price impact dynamics of the partition. Based on the literature review and empirical results it also seems likely that these parameters would be connected directly to scale and type of clustering / HFT activity within the partition. This is a very interesting research direction because the time and company variance and drivers of price impact are not a very well understood subject and at the same time understanding price impact is a crucial problem for many investors.

(2c) What are are the differences in the price impact results based on my data sample compared and the established empirical results in limit order book literature. The results of QPI around trades (figures 10 and 11) are consistent with results of earlier empirical studies (Hautsch and Huang, 2012; Degryse et al., 2005)[4]. Aggregate level features in common with the results are: 1) Market buy (sell) orders move quote prices quote prices up (down) and mirror each other almost perfectly. 2) Market orders to follow a relatively small and decreasing spread and cause an increase in spread even though the both quotes move to the same direction. 3) There is an initial overreaction in the quote where the market order was executed, but it is short lived. 4) After the initial fast reaction, there is a slower and less prominent drift with the same the same direction.

The quantity-quantile-wise split IQPI and LQPI results indicate that a larger the trade quantity is associated with the larger IQPI values. This is very intuitive and is perfectly in line with the findings of e.g. Cont et al. (2013), Hall and Hautsch (2004). It is because the trades will, more often than the smaller, expend the available limit order volume of the best level. The same applies effect also for the LQPI. Unfortunately, the concaveness of the quantity price impact relationship cannot be examined because the quantile split bins are not evenly spaced, but this might be something interesting to examine in the upcoming studies with this data.

Using the data form up to 4 seconds after the event I estimate the permanent absolute quote price impact associated with trades to range between about 0.019% and 0.055% depending on the trade event quantity and side. Similarly, for best level submissions, I get a range between 0.0027% and 0.0071%. Hautsch and Huang (2012) follow a somewhat different method of determining the price impacts and report the results for events where the quantity is 1/2 of the depth on the best level. They find the trades' price impact to be between 0.0205% and 0.0237% and submissions' price impact be between 0.00519% and 0.00557%. Their analysis does not cover cancellations. Although a direct comparison cannot be made, since the

[4]When making comparisons to the Degryse et al. (2005) study it needs to be taken to account that the data sample is quite old (from 1998), likely does not contain a large volume of algorithmic trading and the frequency of trading is much smaller. In fact, the 20 event windows they examine have, depending on the event type, an average clock-time duration of 11-33 minutes. On the other hand, since similarities are none the less found, it indicates that modern AT driven markets function similarly to traditional markets, but simply operate at a much faster pace.

results overlap fully it is plausible that there is no difference between the datasets in this respect. The aggregate QPI effects of best level cancellations are, concerning shape, roughly similar to those of the trades, but smaller in size. The best level submissions, on the other hand, are associated with an impact of opposite direction. All of these results are what would be expected based on the order flow imbalance measure of Cont et al. (2013), as it would treat cancelled, submitted and traded volume exactly the same, considering only if it takes from or adds to the best level.

I find that TPI has jitters at roughly 23 ms interval after the trades. There are no directly comparable results, but it is noteworthy that within each 23 ms period the fastest change towards the long term level where the TPI eventually settles is at the end where the ”response” takes place. Elsewhere the changes will often be smaller or to the other direction. This is consistent with Brogaard et al. (2014) regarding their finding that HFT tend to contribute to the efficiency of price discovery by trading to the direction of permanent price changes. The results on TPI around best level cancellation and submissions particularly interesting because they show an association between the events and immediately after it, a rising cost on market order submitters in realised trades on both sides of the book. While a different kind of empirical test is required for confirmation, this effect could be explained by the finding of Hall and Hautsch (2004) that traders' preference for immediacy appears to increase when the order book reveals a higher dispersion of posted limit prices.

A common result among all of the events IQPI and LQPI results is that the calendar time partitions with weakest event clustering have the strongest effects regarding IQPI and LQPI results. I.e the prices change more when there are less clustered events. This would make sense if clustering were indeed associated with HFT activity as it would be consistent with Brogaard et al. (2014) (US equity market based) findings that HFT are likely to balance liquidity driven short-term price changes.

5. CONCLUSIONS

I give evidence on the evolving order flow activity in OMX Helsinki stock exchange using four measures to study different aspects. Several millisecond environment phenomena in periodicity, clustering, order lifetime and price impact are uncovered.

Regarding periodicity I find that 1) for all of the studied primary events, trades, best level cancellations and best level submissions, there exists an intra-second, 1 ms bin periodicity pattern, that is of comparable scale to the intra-day, 1 minute bin periodicity pattern, persistent over the examined 700 day period when examined in 100 trading day parts and common to studied 21 companies. 2) I find the shape of the intra-second periodicity to be qualitatively similar to the result of Hasbrouck and Saar (2013) but the scale of periodicity is much higher (min-max: 0.9-2.4 vs. 0.96-1.25. In scale where 1 represents uniform distribution).

Regarding clustering 3) I find a response peak that during my sample drifts depending on the event pair at an average rate between -4.8% and -19.0% down to 23 ms. This is in conflict with the results Hasbrouck and Saar (2013) who find that the peak is around 2-3 ms. However, the conflict could be resolved by assuming that in peripheral and small Helsinki markets a lot of order flow comes form London and thus because of network latency (21-22 ms) their processing times would be only 1-2 ms. 4) For all same side event pairs, I find secondary peaks on even intervals such as 1000, 1500 and 2000 ms, which is consistent with order sign autocorrelation results reported by several authors (Degryse et al., 2005; Hautsch and Huang, 2012; Toth et al., 2015). In case of trades and especially the 1500 ms spike it is likely that this is cased by order splitting by a very simple algorithm.

Regarding price impact 5) I find that level of permanent absolute price impact of trades (0.019% - 0.055% depending on event size) and limit orders (0.0027% - 0.0071%) is consistent with earlier results of Hautsch and Huang (2012). 6) Furthermore, I find that in case of each event the quote price impact of bid side event nearly perfectly mirrors the quote price impact of corresponding ask side event which is in line with Zhou (2012).

Based on my results there are several interesting questions worthy of further research. Some of them are connected to the rather strong intra-second periodicity. How does the periodicity affect trading and other metrics of order book and order flow? How should traders plan their actions to take this pattern into account? After all, the intra-second 1 ms bin periodicity in my sample is at a comparable level to the intra-day 1-minute bin periodicity and intra-day periodic effects have been studied a lot and are recognised as something important to take into account.

Another line of questions is related to clustering and the response peak. Is this in fact caused by orders that come from a foreign place (e.g. London) and thus suffer from a rather large network latency? If so how does this kind of periphery market conditions emerge and continue to exist? Can similar effects be found in other Nordic markets? Most importantly, what are the effects on markets where this takes place?

Finally, the third set of questions is related to the order splitting effects: Is it possible to identify the order splitting with e.g. the 1500 ms intervals from the raw order flow with a more sophisticated method revealing the actual sequences of trades. If so what is the price impact of these trades compared to other similar sized trades? I.e. does the market, as expected by the theory, move against the order splitter more than what might be expected based on price impact of single market order?

BIBLIOGRAPHY

ARNUK, S., AND J. SALUZZI (2009): "Latency arbitrage: The real power behind predatory high frequency trading," *Themis Trading white paper*.

BIAIS, B., AND P. WOOLLEY (2011): "High frequency trading," *Manuscript, Toulouse University, IDEI*.

BOUCHAUD, J.-P., J. D. FARMER, AND F. LILLO (2008): "How markets slowly digest changes in supply and demand," *Fabrizio, How Markets Slowly Digest Changes in Supply and Demand (September 11, 2008)*.

BOUCHAUD, J.-P., M. MÉZARD, M. POTTERS, ET AL. (2002): "Statistical properties of stock order books: empirical results and models," *Quantitative Finance*, 2(4), 251–256.

BROGAARD, J., T. HENDERSHOTT, AND R. RIORDAN (2014): "High-frequency trading and price discovery," *Review of Financial Studies*, 27(8), 2267–2306.

CAO, C., O. HANSCH, AND X. WANG (2008): "Order placement strategies in a pure limit order book market," *Journal of Financial Research*, 31(2), 113–140.

CHABOUD, A., B. CHIQUOINE, E. HJALMARSSON, AND C. VEGA (2013): "Rise of the machines: Algorithmic trading in the foreign exchange market," *Journal of Finance, Forthcoming*.

CHLISTALLA, M., B. SPEYER, S. KAISER, AND T. MAYER (2011): "High-frequency trading," *Deutsche Bank Research*, 7.

CONT, R., A. KUKANOV, ET AL. (2012): "Optimal order placement in limit order markets," *arXiv preprint arXiv:1210.1625*.

CONT, R., A. KUKANOV, AND S. STOIKOV (2013): "The price impact of order book events," *Journal of financial econometrics*, 12(1), 47–88.

COPELAND, T. E., AND D. GALAI (1983): "Information effects on the bid-ask spread," *the Journal of Finance*, 38(5), 1457–1469.

DEGRYSE, H., F. DE JONG, M. VAN RAVENSWAAIJ, AND G. WUYTS (2005): "Aggressive orders and the resiliency of a limit order market," *Review of Finance*, 9(2), 201–242.

EISLER, Z., J.-P. BOUCHAUD, AND J. KOCKELKOREN (2012): "The price impact of order book events: market orders, limit orders and cancellations," *Quantitative Finance*, 12(9), 1395–1419.

Ellul, A., C. W. Holden, P. Jain, and R. Jennings (2003): "Determinants of order choice on the New York Stock Exchange," *manuscript, Indiana University.*

Eskelinen, T. (2015): "Order Flow Characteristics In Nordic Stock Exchanges," Master's thesis, Tampere University of Technology, Available: "http://dspace.cc.tut.fi/dpub/bitstream/handle/123456789/23246/eskelinen.pdf ?sequence=1" Referenced 11.04.2016.

Gerig, A. (2015): "High-frequency trading synchronizes prices in financial markets," *Available at SSRN 2173247*.

Gomber, P., B. Arndt, M. Lutat, and T. Uhle (2011): "High-frequency trading," *Available at SSRN 1858626*.

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

Hagströmer, B., and L. Norden (2013): "The diversity of high-frequency traders," *Journal of Financial Markets*, 16(4), 741–770.

Hall, A. D., and N. Hautsch (2004): "A Continuous-Time Measurement of the Buy-Sell Pressure in a Limit Order Book Market," .

Hasbrouck, J., and G. Saar (2013): "Low-latency trading," *Journal of Financial Markets*, 16(4), 646–679.

Hautsch, N., and R. Huang (2012): "The market impact of a limit order," *Journal of Economic Dynamics and Control*, 36(4), 501–522.

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011): "Does algorithmic trading improve liquidity?," *The Journal of Finance*, 66(1), 1–33.

Hollifield, B., R. A. Miller, and P. Sandås (2004): "Empirical analysis of limit order markets," *The Review of Economic Studies*, 71(4), 1027–1063.

Kirilenko, A. A., A. S. Kyle, M. Samadi, and T. Tuzun (2011): "The flash crash: The impact of high frequency trading on an electronic market," *Available at SSRN 1686004*.

Linnainmaa, J. T. (2010): "Do limit orders alter inferences about investor performance and behavior?," *The Journal of Finance*, 65(4), 1473–1506.

Lo, I., and S. G. Sapp (2010): "Order aggressiveness and quantity: How are they determined in a limit order market?," *Journal of International Financial Markets, Institutions and Money*, 20(3), 213–237.

MASLOV, S., AND M. MILLS (2001): "Price fluctuations from the order book perspective—empirical facts and a simple model," *Physica A: Statistical Mechanics and its Applications*, 299(1), 234–246.

NASDAQ (2010): "NASDAQ ICTH feed documentation," `http://www.nasdaqtrader.com/content/technicalsupport/specifications/dataproducts/tv-itch2.pdf`, Accessed: 18.07.2016.

PARLOUR, C. A., AND D. J. SEPPI (2008): "Limit order markets: A survey," *Handbook of financial intermediation and banking*, 5, 63–95.

RANALDO, A. (2004): "Order aggressiveness in limit order book markets," *Journal of Financial Markets*, 7(1), 53–74.

SIIKANEN, M., J. KANNIAINEN, J. VALLI, AND A. LUOMA (2016): "Limit Order Books and Liquidity around Scheduled and Non-Scheduled Announcements: Empirical Evidence from NASDAQ Nordic," *Manuscript. Available at SSRN 2547461.*

TOIVONEN, V.-M. (2013): "High frequency price impact of order book events-Evidence from Nordic equity markets," Master's thesis, Aalto University School of Economics, Department of Finance, Available: http://urn.fi/URN:NBN:fi:aalto-201305102107 Referenced 20.3.2016.

TOKE, I. M. (2011): ""Market making" in an order book model and its impact on the spread," in *Econophysics of Order-driven Markets*, pp. 49–64. Springer.

TOTH, B., Y. LEMPERIERE, C. DEREMBLE, J. DE LATAILLADE, J. KOCKELKOREN, AND J.-P. BOUCHAUD (2011): "Anomalous price impact and the critical nature of liquidity in financial markets," *Physical Review X*, 1(2), 021006.

TOTH, B., I. PALIT, F. LILLO, AND J. D. FARMER (2015): "Why is equity order flow so persistent?," *Journal of Economic Dynamics and Control*, 51, 218–239.

TUOMINEN, A. (2012): "The role of High Frequency trading in limit order book activity: Evidence from Helsinki Stock Exchange," Master's thesis, Aalto University School of Economics, Available: https://aaltodoc.aalto.fi/bitstream/handle/123456789/3562/hse_ethesis_12823.pdf?sequence=1&isAllowed=y Referenced 20.7.2016.

ZHAO, L. (2010): "A model of limit-order book dynamics and a consistent estimation procedure," Ph.D. thesis, Citeseer.

ZHOU, W.-X. (2012): "Universal price impact functions of individual trades in an order-driven market," *Quantitative Finance*, 12(8), 1253–1263.