HEC MONTRÉAL
École affiliée à l'Université de Montréal

High-frequency data : Information processing and financial analysis

par
Yann Bilodeau

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

Décembre 2021

HEC MONTRÉAL
École affiliée à l'Université de Montréal

Cette thèse intitulée :

High-frequency data : Information processing and financial analysis

Présentée par :

Yann Bilodeau

a été évaluée par un jury composé des personnes suivantes :

Tolga Cenesizoglu
HEC Montréal
Président-rapporteur

Georges Dionne
HEC Montréal
Directeur de recherche

Vincent Grégoire
HEC Montréal
Membre du jury

Anatoliy Swishchuk
University of Calgary
Examinateur externe

Iwan Meier
HEC Montréal
Représentant du directeur de HEC Montréal

Résumé

Dans un format classique, cette thèse porte sur l'utilisation de données financières à haute fréquence. Nous travaillons sur un ensemble de données provenant du système Xetra Enhanced Broadcast Solution (EnBS) couvrant la période du 1er février au 30 avril 2013. Cet ensemble de données contient toutes les informations qui ont été fournies par Deutsche Börse à ses clients payant en temps réel, qu'ils soient des banques, des investisseurs institutionnels ou des fonds de gestion. Bien que très riches en informations sur le carnet d'ordres limites et sur les transactions exécutées, en raison de la complexité de leur format initial, ces données ne sont pas directement utilisables à des fins de recherche en finance. Ainsi, les trois premiers chapitres de cette thèse consistent en l'élaboration d'une méthodologie incrémentale permettant de traiter ces données et d'extraire et déduire des informations de plus en plus complexes et détaillées. Dans les chapitres quatre et cinq, nous utilisons les données exclusives ainsi produites pour travailler sur des réelles questions financières.

Au Chapitre 1, nous élaborons un processus permettant d'obtenir l'état exact du carnet d'ordres pour une action donnée et ce, à n'importe quel moment d'une journée de transaction. Cette méthodologie produit des séries chronologiques très simples qui deviennent facilement utilisables à des fins de recherche. Cependant, étant donné qu'elles sont dérivées de notre ensemble de données Xetra EnBS, elles ne fournissent aucune information concernant les ordres limites réels. En effet, l'information de carnet d'ordres fournie par Xetra consiste uniquement en des mises à jour du nombre d'ordres et de la quantité totale d'actions disponible à chaque niveau de prix. Par conséquent, au Chapitre 2, nous développons une seconde méthodologie qui permet d'utiliser ces variations afin d'identifier et de caractériser les événements de carnet d'ordres qui sont à l'origine des données que nous observons. Nous identifions les événements les plus fréquents reliés à l'addition, la suppression et l'exécution de la liquidité. Ces évènements étant fort probablement liés à la soumission, l'annulation et l'exécution des ordres limites, nous nous retrouvons avec une description très complète de l'activité boursière pour chacun des titres d'intérêt. En ce qui concerne les titres composant les indices DAX, MDAX et SDAX

nous considérons identifier avec succès 97,4 %, 98,6 % et 99,6 % des événements de carnet d'ordres.

Au Chapitre 3, nous utilisons les règles d'identification d'événements élaborées au Chapitre 2 comme point de départ pour le développement d'une troisième méthodologie qui nous permet le suivi des ordres limites. Grâce à cette deuxième extension de notre méthodologie initiale, nous sommes en mesure de créer et de maintenir une liste d'ordres limites pour chaque niveau de prix, bien qu'aucune information en ce sens ne soit initialement fournie dans les données Xetra. Pour les composantes des indices DAX, MDAX et SDAX, en moyenne, nous considérons pouvoir suivre avec succès 41,5%, 87,4% et 98,6% des ordres, et ce, à partir de leur soumission jusqu'à une situation où elles sont annulées, totalement exécutées ou qu'elles sortent de la partie visible du carnet d'ordres. Nous observons une relation log-linéaire négative entre le nombre moyen pondéré par le temps du nombre d'ordres se trouvant sur les cinq premiers niveaux de prix et nos taux de succès.

Au Chapitre 4, nous analysons la structure de dépendance de l'arrivée des événements du carnet d'ordres sur la base de processus de Hawkes multivariés. En utilisant les résultats produits au Chapitre 2 comme source, nous trouvons des relations récurrentes entre des événements appartenant à 86 catégories différentes. En plus des transactions, ce groupe d'évènements inclue les soumissions d'ordres limites et les annulations ayant lieu jusqu'au vingtième niveau de prix du carnet d'ordres. Nous nous concentrons sur BMW, SAP et ADS, trois titres très liquides de l'indice DAX. Pour chacun de ces titres, nous construisons un modèle descriptif en sélectionnant les relations les plus récurrentes. Estimés pour chaque journée de transaction, nous constatons que les modèles ainsi définis offrent des performances très intéressantes, en particulier pour les soumissions d'ordres limites et les annulations survenant sur les cinq premiers niveaux de prix du carnet d'ordres. En utilisant les nombreux paramètres estimés, nous sommes en mesure de décrire une dynamique que nous pouvons relier aux comportements des nombreux intervenants de marché qui peuvent avoir des objectifs très différents.

Finalement, au Chapitre 5, nous analysons divers phénomènes à partir des caractéristiques des ordres limites. En travaillant sur les titres ayant présenté les taux de réussite en termes de suivi des ordres au chapitre 3, il devient possible de se concentrer sur différents aspects relatifs à ces véhicules de liquidité. En mettant l'emphase sur le niveau de prix d'arrivée et le temps écoulé entre la soumission d'un ordre et son annulation, nous constatons qu'une proportion importante des ordres est soumise et annulée sans réel potentiel d'exécution. Par la suite, nous travaillons sur les ordres qui semblent soumis et annulés en séquences. Nous relions ces opérations à des activités potentielles de *quote-stuffing*. Nous développons et mettons en œuvre une méthodologie permettant de regrouper ces séquences sur la base de leurs caractéristiques globales et de celles des ordres qui les composent. Ainsi, nous sommes en mesure d'identifier et de qualifier les algorithmes que nous considérons responsables de ces activités.

Mots clés : Carnet d'ordres limites, microstructure des marchés financiers, processus de Hawkes, estimation par maximum de vraisemblance, liquidité, *quote-stuffing*.

Méthodes de recherche : Recherche quantitative.

Abstract

Presented in a classic format, this thesis focuses on high-frequency financial data usage. We work on a Xetra Enhanced Broadcast Solution (EnBS) historical dataset covering the time period going from February 1st to April 30th, 2013. This dataset contains all the information that have been provided by Deutsche Börse to their Frankfurt Stock Exchange Xetra real-time paying customers. Although very rich in limit order book (LOB) and executed transactions information, because of its initial format complexity, the data is not directly usable for financial research purposes. Therefore, the first three chapters consist in the development of an incremental methodology allowing to process Xetra data in order to extract and deduce increasingly complex and detailed information. In chapters four and five, we use the exclusive resulting data to work on high-frequency financial questions.

In Chapter 1, we present a process allowing to obtain the exact order book state for a given stock at any given time during a trading day. This methodology generates time series that are easily usable for research purposes. However, we have to keep in mind that since they are tailored to our Xetra EnBS dataset, they provide no information regarding the actual limit orders. Indeed, the Xetra order books information is essentially diffused through updates in the number of standing orders and total quantity of shares. Consequently, in Chapter 2, we develop a methodology allowing to use these variations in order to identify and characterize the order book events. We identify the most frequent events as liquidity added, liquidity removed and liquidity executed, which are potentially related to actual limit orders submission, cancellation, and execution. We consider that we successfully identify 97.4%, 98.6% and 99.6% of the events affecting the order books of the DAX, MDAX and SDAX indexes components.

In Chapter 3, we use the Chapter 2 events identification rules as entry points to develop a limit order tracking methodology. Through this second extension, we present a way to build and maintain a limit order list for each visible price level. For the DAX, MDAX and SDAX indexes components, we consider being able to successfully track an average of 41.5%, 87.4% and 98.6% of the orders from their submission to a situation where they are cancelled, totally executed or they exit the visible part of the book. We find a negative

log-linear relationship between the time-weighted average number of orders standing on the first five levels of the book and our success rates.

In Chapter 4, we analyze the limit order book events occurrence dependency structure by defining high-dimensional Hawkes processes models. Using Chapter 2 results as a data source, we seek for recurrent relationships among events from a set of 86 event types which in addition to transactions, includes limit order submissions and cancellations taking place up to the 20th depth level of the order book. We focus on BMW, SAP, and ADS, three liquid DAX index stocks. For each of these stocks, we build a tailored descriptive model by selecting the most recurrent events relationships. Estimated on a daily basis, we find that the selected models offer interesting data fitting performance, particularly for limit order submissions and cancellations occurring on the first five price levels of the order book. Using the comprehensive sets of estimated parameters, we describe a global events arrival dynamics that we relate to the potential behaviors of market participants having different objectives and directional views.

Finally, in Chapter 5, we analyze different market phenomenon from the point of view of the limit orders characteristics. Indeed, by working on the stocks having presented the highest tracking success rates in Chapter 3, it becomes possible to focus on different aspects of these elements that actually represent liquidity vehicle on the Xetra stock market. Then, we combine the price levels of arrival and the time elapsed between an order submission and its cancellation to find that an important proportion of limits orders are actually submitted and cancelled with virtually no execution potential. Second, we work on the limit orders that may have been submitted and cancelled in sequences. We relate these operations to potential quote-stuffing activities. We develop and implement a methodology allowing to group theses orders sequences on the basis of their global characteristics and those of the orders identified as part of them. In this way, we identify and qualify the algorithms that we consider responsible for these order sequences.

Keywords : Limit order book, market microstructure, Hawkes processes, maximum-likelihood estimation, liquidity, quote-stuffing.

Research methods : Quantitative research.

Table of contents

List of tables

List of figures

List of acronyms

- **BPI** Best Price Impact
- **HFT** High Frequency Trading
- **LO** Limit Order
- **LOB** Limit Order Book
- **LOC** Limit Order Cancellation
- **LOS** Limit Order Submission

À Martine, l'amour de ma vie et à ma fille Olivia, qui n'a rien connu d'autre que le doctorat depuis sa naissance. Je vous promets d'essayer de faire en sorte que la vie soit maintenant plus normale pour notre famille.

Remerciements

Je tiens d'abord à remercier le professeur Georges Dionne de m'avoir donné l'opportunité de réaliser mon doctorat à l'intérieur de la Chaire de Recherche du Canada en Gestion des risques. Je connaissais son excellente réputation avant d'entreprendre ces études et je tiens à spécifier que je n'ai jamais été déçu. Je le remercie pour son support, ses conseils, son soutien financier, mais aussi pour sa patience à travers toutes ces années. L'avant-gardisme et la vision du professeur Dionne a fait en sorte que nous avons été parmi les premiers à avoir la chance d'apporter une contribution dans certains domaines de pointe, ce qui a permis l'obtention de résultats très intéressants par chacun des membres de l'équipe. Sur ce point, je remercie mes collègues Gabriel Yergeau, Xiao-Zhou Zhou et Mathieu T-Blais pour leurs commentaires et leurs excellentes idées. Grâce aux réunions organisées par le professeur Dionne, nous avons été en mesure d'intégrer une notion d'équipe dans des travaux qui se veulent habituellement de nature très individuelle.

D'un point de vue plus personnel, je tiens à remercier ma mère Huguette, qui m'a aidé à passer au travers grâce à ses prières et ses appels presque quotidiens. Dans cette lignée, je remercie également ma belle-mère Nicole et mon beau-père Yvan, deux autres adeptes de la prière qui se sont avérés très importants dans mon parcours. Je tiens aussi à remercier mon père Yvon et ma sœur Marjorie. Dans le contexte d'une thèse de doctorat, il arrive parfois que le meilleur support que l'on puisse offrir soit de ne pas poser trop de questions sur l'avancement des choses et ils l'ont très bien compris. Finalement, je remercie mes meilleurs amis Martin, Dany et Jacques pour m'avoir aidé à traverser ces années parfois difficiles avec le sourire.

C'est avec beaucoup de reconnaissance que je souligne également le support financier de la Chaire de Recherche du Canada en Gestion des risques, du Fonds de recherche du Québec – Société et culture (FRQSC), de l'Institut canadien des dérivés, du Centre interuniversitaire sur le risque, les politiques économiques et l'emploi (CIRPÉE) ainsi que du Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport (CIRRELT).

Preface

Over the last decades, data exploitation has become a key factor of success in most, if not all business areas. The financial sector has been widely affected by this trend, especially when it comes to stock markets. Indeed, with the proliferation of algorithmic and high-frequency trading, the amount of available data regarding the activity affecting a simple stock has exploded. Even though data processing and exploitation do not represent a pure financial matter, this topic, which essentially combines high-performance computer and data science, is now essential to produce high-level academic and professional work in Finance. In this context, we make it the backbone of this thesis. We attempt to demonstrate that smart data exploitation provides an advantage in financial research.

We dedicate the first three chapters to the formalization, processing, and manipulation of an important dataset obtained from Xetra, the electronic platform of the Frankfurt Stock Exchange. From one of these chapters to another, we develop an incremental methodology allowing to obtain more detailed and relevant information from our original dataset, particularly regarding the activity taking place on the limit order book of the stocks listed on this market. Keeping in mind that this thesis is actually presented in the context of a doctoral candidacy in Finance, in chapters four and five we use the produced data to work on pure financial issues.

Introduction

Frankfurt Stock Exchange operations are now mostly conducted on Xetra, its electronic trading venue. This platform disseminates executed trades, limit order book (LOB) content and other market information through software interfaces like Xetra Enhanced Broadcast Solution (EnBS). Market participants such as banks, fund managers, hedge-funds and algorithmic traders have to subscribe to this type of paid service to access real-time Xetra data. Deutsche Börse, the company responsible for Xetra operation and maintenance, provides its customers with the technical specifications required to develop their own data consuming applications. In order to utilize the market information, such applications have to establish connections to Xetra real-time data feeds and decode the structures used as information vehicles. Besides real-time data, Deutsche Börse also sells Xetra historical datasets. It is then possible for market outsiders, as us, to access integral market data that has once been broadcasted to online subscribers. In this context, the Canada Research Chair in Risk Management acquired a Xetra EnBS dataset covering the time period going from February 1 to April 30, 2013.

From an academic point of view, with its microsecond (10^{-6} second) time precision, our Xetra dataset represents a very rich source of ultra-high frequency financial information. However, such data are almost impossible to use in their initial shape for research purposes because of their complex format designed for real-time broadcasting efficiency. To overcome this issue, in Chapter 1, we present a complete methodology allowing to make Xetra data usable in a research context. By formalizing the original data, we are able to develop a complete LOB states management system. Once implemented, it becomes possible to obtain the exact LOB state for a stock at any time on a given trading day and this, in an easily usable form. The results from this first chapter methodology have already been used as data sources for different theses and papers in the context of the Canada Research Chair in Risk Management activities.

While important segments of the financial research on intraday and high-frequency stock markets activities involve modeling the arrival of orders submission, cancellation, and execution, the Xetra EnBS data does not provide direct information regarding these events. Consequently, in Chapter 2, we extend the concepts defined in Chapter 1 and develop a methodology allowing to recover the order book events occurrence information. This task presents several challenges since providing this type of information does not seem to be a part of the Xetra EnBS objectives. Indeed, the LOB information is aggregated and provided in an incremental form. Therefore, in order to detect and characterize the occurrence of liquidity events that we qualify as liquidity added, liquidity removed, and liquidity executed, we have to use the variations in the number of standing orders and total quantity of shares as well as trades information. Unlike the methodology presented in Chapter 1, which is exclusive to the Xetra EnBS data structure, our flexible events identification process could be adapted and implemented on any price level aggregated diffusion system.

Chapter 3 is the last one mostly focussing on Xetra data processing. In a second extension of our Chapter 1 methodology, this chapter goes way further in retrieving information that is not directly provided in our Xetra EnBS dataset. It focusses on the limit orders themselves. As claimed before, such dataset provides no information regarding the orders that actually compose the book. However, having identified the liquidity added event as the entry point of an order in the book and the liquidity removed or liquidity executed events as potential exit points, we consider possible to relate these elements and track an order over its complete life cycle. In order to produce highly accurate information, we build a methodology allowing for a minimum of assumptions and hypothesis. We acknowledge that this zero assumption philosophy may result in less orders considered as successfully tracked from their submission to their cancellation or execution, However, we believe that the orders that we finally consider as successfully tracked have a high potential of matching their underlying counterparts actually submitted, cancelled, and executed on Xetra. In fact, for the sake of intellectual integrity, we choose to end up with a dataset including fewer orders for which we are almost certain of the outcome rather than one that would also contain less reliable information.

It is in Chapter 4 that we really enter the financial core of this thesis. In this chapter, we mainly focus on the arrival of order book events and their interdependence. We model these elements using multivariate Hawkes processes. Similar works have already been done in the past, but with a more limited scope than the analysis proposed in this chapter. Indeed, these works has generally focused on the analysis of the arrival of transactions and events affecting the first levels of the order book. However, given the event identification methodology developed in Chapter 2 and the richness of our Xetra dataset, we are able to obtain a much more important set of events than what has been studied before. In fact, for both the bid and the ask sides of the order book, in addition to transactions, we have the information regarding order submissions and cancellations taking place up to the twentieth price levels. Assuming that each of these events may be part of the global order book dynamic, we end up with 86 potentially interrelated events types. For generality and inclusiveness purposes, we undertake this work with the initial assumption that most of these events are potentially interrelated. Thus, as a first step, we develop and implement an estimation methodology that allows us to identify the most recurrent event relationships on a daily basis, for a stock of interest. We apply this methodology to the events identified for BMW, ADS and SAP, three blue chips traded on Xetra during the 61-trading days period going from February 1 to April 30, 2013. Once the recurrent relationships have been identified, we are able to qualify and quantify them using the estimated Hawkes processes parameters. Therefore, in a second step, we undertake a general but exhaustive analysis and interpretation task regarding all these relationships, with the goal of relating them to the potential behaviors of different market participants whether they are buyers, sellers, patient, impatient, liquidity provider, liquidity consumer and so on. Although this work is laudable, we have to keep in mind that for each of our three stocks, it essentially represents an attempt to describe the entire order book dynamics using a single model and aggregated parameters set for the 61 trading days. Nevertheless, we believe that we are able to achieve this task by identifying several interesting trends and phenomena with regard to these dynamics.

Finally, in Chapter 5, we focus on the limit orders themselves, most particularly those ending with a cancellation. In this context, we work on the successfully tracked orders obtained using the order tracking methodology developed in Chapter 3. First, we analyze the orders duration, which we define as the time elapsed between an order submission and its cancellation. Using this information, we are able to identify the orders with a high potential of having been submitted in the context of algorithmic or high-frequency trading activity. Mostly focusing on these specific orders, we analyze their journey in the limit order book in terms of position changes with regard to their price level of submission. Then, we focus on the actual execution possibility of these orders with regard to their duration and price level of submission. In the second part of this chapter, we study the possibility that short duration orders may be submitted in sequences, in operations that can take the appearance of quote-stuffing activities. Therefore, we develop a simple methodology allowing us to identify short duration orders appearing to be part of the same sequence. Afterward, we focus on orders sequence that we identify with a high potential of being the result of quote-stuffing activities. By considering the characteristics of the sequences themselves and those of the orders that compose them, we develop a frame of reference that allows us to identify, classify and group together the algorithms that may be at the source of these sequences. We are therefore able to present several examples of behavior of algorithms that seem to be involved in quote-stuffing operations. Finally, after a review of the German regulations regarding algorithmic and high-frequency trading activities that entered into force within the year following our data period, we perform a final general analysis focusing on the liquidity provided with respect to the limit orders duration.

Chapter 1
Xetra data engineering

Different types of market information can be retrieved from Xetra EnBS historical data. However, some of the most interesting elements are unusable in the format trough which they are initially provided. To overcome this issue, we have developed XetraParser, a piece of software designed to improve the convenience of Xetra EnBS historical data. It is central to our most basic data processing cycle. Because it is more technical than financial, this infrastructure is described in Appendix 1. All tables and figures are at the end of the chapter.

Of all the Xetra available information, the executed transactions represent the simplest elements to obtain. No complex mapping task is actually required to get the traded number of shares and their execution prices. They are broadcasted using a simple data structure grouping all the relevant information. Once formatted, they become the simplest XetraParser output. Auction data retrieval is similarly relatively straightforward. Once processed, the opening, closing and mid-day auctions information also become a simple output. We format it to make the potential matching price, matching quantity, and surplus quantity easily available for any stock. On the other hand, we can easily state that the information related to the limit order book represents the most challenging part of a Xetra Enhanced Broadcast Solution (EnBS) historical dataset. It is actually impossible to obtain a tick-by-tick LOB information by simply observing the dataset since its content is communicated through an incremental information system. Getting the exact state of an asset LOB for a given time generally involves maintaining an internal copy of the book and updating it using the provided increments. In this context, based on our XetraParser output, we have first developed an order book rebuilding component. However, even with this tool on hand, the issue of producing data that can be used for research purposes remains intact. Consequently, the main purpose of this chapter is to presents LOB state tracking methodology that may be used to produce datasets that are well suited for financial research.

This chapter is organized in the following way : Section 1.1 presents the Xetra EnBS information required by our LOB rebuilding and liquidity tracking tasks. Still based on the Xetra model, Section 1.2 develops a basic order book state model. Section 1.4 extends the model and describes our order book state tracking methodology. Section 1.5 presents a second model extension and develops our liquidity events identification methodology. Additionally, Appendix 1.1 presents our Xetra data processing cycle.

1.1 Xetra EnBS data structure

In order to broadcast market information, Xetra EnBS uses a system of messages in which each message type carries its own category of data. Xetra EnBS offline datasets initially take the shape of large data dumps consisting of all messages produced and sent to online clients over a given time period. As an online service subscriber would do, we must decode these messages in order to access their informational content. Fortunately, the task is simpler for us than for these clients since we already have a physical copy of the messages in hand, which means that we do not have to care about correctly getting them through real-time network communications. Out of the eighteen different message types, we identify only four as responsible for actual financial information transmission. The other types of message are related to a wide range of technical aspects arising from the system and its operation. We discard this technical data and focus on the financial messages needed for LOB rebuilding and trades extraction.

In a first step, we use *Instrument Reference Data* messages to extract information on the assets traded on the Xetra platform. They provide general info such as the instruments internal identifier, long name, symbol, and group. We also use these messages to get each asset *market depth*, which corresponds to the maximum number of visible price levels for each side of the LOB. Depending on the asset liquidity, five, ten or twenty levels may be visible. For the stocks considered in this chapter, which correspond to the DAX, MDAX and SDAX indexes components, twenty price levels are made available for each side of the book.

LOB incremental information is distributed through *Inside Market Delta* (*delta*) messages. From our point of view, they represent the heart of the so-called ultra-high frequency dataset. Indeed, all visible LOB state change are reported through delta messages. Obviously, since those changes occur any time, often several times in a single trading second, these messages are not evenly distributed in time. This fact makes the *delta* message microsecond time stamps very interesting. On the other hand, *Inside Market Snapshot (snapshot)* messages periodically provide a complete state of the twenty visible price levels for both sides of the book. Since it is possible to track the LOB state from the beginning of a trading day by only using delta messages, in normal conditions, we do not use *snapshot* messages for rebuilding purposes. However, we use them to ensure the validity of our internal order books representation. Each time a *snapshot* message is encountered, its content is compared with the current state of our LOB in order to detect eventual mismatches. We also use them as new starting points in cases where one or more *delta* messages are missing. These infrequent situations are identified using the sequence number included in each delta *message*. Missing *delta* messages are typically concentrated on short time periods and appear to simultaneously affect multiple assets. This suggests that they result from Xetra EnBS system failures. Inverted messages also represent an issue observable in the dataset. We easily manage this issue by buffering multiple messages and ordering them correctly before their treatment.

All executed transactions are diffused through *All Trade Price* (*ATP*) messages. Those having a visual effect on the LOB are also reported in a subsection of the delta message broadcasting the related order book updates. In such cases, we assume the *delta* message timestamp to correspond to the time from which the order book is affected by the trade. On the other hand, some trades or trade segments do affect the LOB appearance. For examples, the execution of some *iceberg* or *hidden* orders leaves the book visually intact. Consequently, these trades are only reported in *ATP* messages. For this reason, we rely on both *delta* and *ATP* messages to get all executed transactions information.

When diffusing one or more LOB state changes, a delta message contains a subsection providing all the information needed by subscribers to update their own LOB representations. Each line of this subsection corresponds to an update *instruction* (*or action*) that must be applied to LOB for it to remain up-to-date and consistent. We count five types of these instructions. Depending on their characteristics, they can affect up to twenty price levels. First, assuming BS to represent the LOB sides set $\{bid, ask\}$, we use the following expression to describe the set of all possible *new element* instructions :

(1.1)	$NEW = \{(t, s, l, p, n, q) \in \mathbb{R}^+ \times BS \times \mathbb{N}^+ \times \mathbb{R}^+ \times \mathbb{N}^+ \times \mathbb{N}^+ \mid l \leq 20\}.$

Therefore, a *new element* action $(t^n, s^n, l^n, p^n, n^n, q^n) \in NEW$ involves the insertion of a new price level at position l^n of the book side s^n. It represents q^n newly visible shares distributed among n^n limit orders for which the price is p^n. Also, we consider t^n, which correspond to the global delta message timestamp, as the time from which the change is fully integrated to the LOB. Instructions of this type are observed in two main situations. First, they are sent in order to fill up an empty order book at the beginning of each continuous trading period. Such periods begin after the opening and intraday auctions or, in more tumultuous situations, after a volatility interruption. In these cases, twenty new elements instructions are encountered for each LOB side. In the second situation, which is the most common, the *new element* instruction is used to signal the creation of a new price level, at position l^n, inside a non-empty LOB. Existing elements for which the depth level number is greater or equal to l^n are then moved up by one position as shown in the Table 1.1 first example.

Since Xetra EnBS allows only twenty visible price levels, a new order book element inserted in an already full LOB side causes the element with the highest level number to be excluded from this *visibility window*. We have to keep in mind that in such a situation, although becoming invisible, the excluded element probably still exists[1].

In a similar way, the following expression introduces the set of all possible *element change* instructions, which corresponds to the second, and simplest LOB update action :

(1.2)	$CHANGE = \{(t, s, l, n, q, v) \in \mathbb{R}^+ \times BS \times \mathbb{N}^+ \times \mathbb{N}^+ \times \mathbb{N}^+ \times \mathbb{N} \mid l \leq 20\}.$

A *change element* instruction $(t^c, s^c, l^c, n^c, q^c, v^c) \in CHANGE$ indicates a modification taking place on the price level number l^c of the LOB side s^c. The modified price level

[1] Based on the Xetra EnBS specifications, we claim that no assumption should be made about price levels that are not in the visible part of the book since their consistency cannot be certified.

number of orders and aggregated quantity of shares are provided by n^c and q^c. They represent the only values that can actually be updated using this instruction category. As claimed before, it is possible for the price level modification to be reported concurrently to the execution of one or more shares with a corresponding price. In this case, we assume this quantity to be reported through $v^c > 0$. This LOB update action type does not affect the other price levels. The second example presented in Table 1.1 corresponds to an *element change* instruction.

Using the same notation, expression (1.3) covers the *delete element* LOB update action set.

(1.3)	$DELETE = \{(t, s, l, v) \in \mathbb{R}^+ \times BS \times \mathbb{N}^+ \times \mathbb{N} \mid l \leq 20\}.$

In this third case, for any $(t^d, s^d, l^d, v^d) \in DELETE$, the price level number l^d of the LOB side s^d is removed. Once again, we would consider a simultaneous traded volume whose price would correspond to that of the deleted level to be reported through $v^d > 0$. As a side effect of a level deletion, the undeleted price levels for which the number is larger than l^d move down by one position, which is observed in the third example of Table 1.1. Most of the time, deleted elements are offset by an equivalent number of *new element* instructions in order to maintain twenty visible price levels in the LOB. We claim this added liquidity to generally correspond to price levels already existing outside the visibility window.

Xetra EnBS also includes the *delete through* and *delete from* update instructions. Similar to delete element action, they share the same set of possible values. In the case of a delete through instruction $(t^{dt}, s^{dt}, l^{dt}) \in DELETE$, the side s^{dt} price levels 1 to l^{dt} are deleted, causing all the remaining price levels to upgrade by l^{dt} elements in the execution priority structure. It is used to signal the removal of more than one price level in a single instruction. From an information broadcasting point of view, this way is more efficient than using multiple *delete element* actions. Nevertheless, in order to simplify our own procedures, we replace each *delete through* update action by l^{dt} simple *delete element* instructions. To minimize the impact of this substitution, we place these instructions in a descending order, sequentially affecting levels l^{dt} to 1.

Finally, the *delete from* instruction $(t^{df}, s^{df}, l^{df}) \in DELETE$ signals the removal of the side s^{df} price levels numbers l^{df} and over. Appearing as an effective way to signal an empty LOB, a *delete from* action for which $l^{df} = 1$ is generally submitted for both sides of the book just before the intraday and the end-of-day auctions. As before, we replace *delete from* instructions by a set of simple *delete element* actions.

1.2 The limit order book

In this section, we define the foundations of the LOB model that will be extended to achieve our states, events and orders tracking goal. While adaptable, this model is fundamentally designed to work well with our Xetra data source. For example, it fully supports the concept of numbered price levels. This feature being omnipresent in the Xetra EnBS LOB system, its integration makes it easier to link with the data. Thus, this segment focuses on describing sets and functions representing LOB properties and mechanics. Given the nature of the work done in this chapter, we define a totally deterministic LOB model. From our point of view, characterizing price levels states, detecting liquidity flow events and tracking limit orders as far as possible in their life cycle is better achieved in a non-stochastic environment.

We chose to present the evolution of our LOB on a state basis rather than time. Even though time characterizes all the objects we attempt to track, we do not use it as the LOB evolution scale. In our Xetra EnBS context, even with a microsecond precision, multiple order book updates can be concurrently reported. Although correctly ordered for LOB rebuilding purposes, it is often impossible to establish their actual occurrence sequence. To overcome this issue, we assume the LOB to evolve from state to state regardless of the elapsed time. Each time a price level change has happened, we consider the whole LOB to enter a new state. Using a notation borrowed from the measure theory, we define $\Omega = \{\omega_i : i \in \mathcal{I}\}$ and $\mathcal{I} = \{1,2,3, \ldots\}$ to represent the set of all states and their indexes[2]. In order

[2] Geneviève Gauthier, 2007, Acétates Introduction à la probabilité.

to make the notation lighter, concerned elements are presented with subscript parameters i or $i+1$ referring to consecutive states ω_i and ω_{i+1}. These two states are sufficient since all our identification tasks are defined using a single LOB update. Moreover, we consider ω_1 to represent the first LOB state of a Xetra continuous trading period. In addition to the beginning of the trading day, those states can occur in any situation where the continuous trading resumes after a suspension, which can result from an auction, a volatility interruption, and so on.

Based on the previously defined book sides set BS, we define the function $L_i: s \in BS \rightarrow \mathbb{N}$. $L_i(s)$ provides the book side s highest visible price level number on state $\omega_i \in \Omega$. Using these notions, the following expression presents our price function :

(1.4)	$p_i: \{(s,l) \in BS \times \mathbb{N}\} \rightarrow \mathbb{R}^+$ $p_i(s,l) = \begin{cases} (s,l)\ price, & L_i(s) > 0 \wedge l \leq L_i(s) \\ 0, & L_i(s) = 0 \vee l > L_i(s) \end{cases}$

For any state $\omega_i \in \Omega$, this function relates a LOB side and price level number combinations (s,l) to an actual price value. We do not attempt to provide a formula for this relation since it is observed rather that computed. However, we establish that $p_i(s,l)$ only provides a value for actual visible price levels, which is enforced using $L_i(s)$. It maps any other natural number to the empty set. We consider this function as the link between the Xetra EnBS numbered price levels concept and more traditional LOB representations based on the price.

In any state $\omega_i \in \Omega$ such that $L_i(bid) > 0$ and $L_i(ask) > 0$, $p_i(bid,1)$ and $p_i(ask,1)$ represent the best bid and ask prices. In the same situations, we consider the LOB price structure to be consistent when the following conditions are satisfied :

(1.5a)	$\forall\, l_1, l_2 \in \mathbb{N} \,:\, 0 < l_1 < l_2 \,\leq L_i(bid),\;\; p_i(bid, l_1) > p_i(bid, l_2)$
(1.5b)	$\forall\, l_1, l_2 \in \mathbb{N} \,:\, 0 < l_1 < l_2 \leq L_i(ask),\;\; p_i(ask, l_1) < p_i(ask, l_2)$
(1.5c)	$\forall\, l_a, l_b \in \mathbb{N} :\, 0 < l_a \leq L_i(ask) \wedge\, 0 < l_b \leq L_i(bid),$ $p_i(ask, l_a) > p_i(bid, l_b)$

As we can see, (1.5a) and (1.5b) ensure the price to be a decreasing (increasing) function of the bid (ask) side level number. Condition (1.5c) certifies that bid prices are smaller than any ask price.

Finally, we define the following functions in order to map a price level whose price is p to its total number of visible orders and aggregated quantity of shares.

(1.6)	$n_i : p \in \mathbb{R}^+ \rightarrow \mathbb{N}$
(1.7)	$q_i : p \in \mathbb{R}^+ \rightarrow \mathbb{N}$

Once again, for a given state $\omega_i \in \Omega$, $n_i(p)$ and $q_i(p)$ simply report the observed relations between two sets of values. However, we consider these functions to return zero for price values that are not related to actual Xetra visible price levels. Thus, for $\forall p \in \{p \in \mathbb{R}^+ \big| (\forall s \in BS)(\forall l \in \mathbb{N}^+)[p_i(s, l) \neq\, p]\}$, $n_i(p) = 0$ and $q_i(p) = 0$.

Also, we assume the following price level consistency condition to be satisfied in any state $\omega_i \in \Omega$:

(1.8)	$\forall\, p \in \mathbb{R}^+, q_i(p) \geq n_i(p)$

This expression states that a limit order must contain a minimum of one share. Although appearing obvious, it is essential to the definition of our LOB events and limit orders tracking procedures.

In general terms, since Xetra EnBS refers to liquidity using a combination of order book side and price level number, we define the function $p_i(s, l)$ to translate this information into a simple $p \in \mathbb{R}^+$ value that can be used as an argument for $n_i(p)$ and $q_i(p)$.

1.3 Order book states

In this section, we slightly extend our model in order to implement a LOB state tracking methodology. The main purpose of this procedure is to transform Xetra EnBS LOB incremental information into a dataset from which it is straightforward to get the exact state of the book at any time. For both logical and technical reasons, we choose to model our LOB state system on a price level basis. Indeed, it would be resource consuming to replicate the complete order book for each price level state change. Therefore, whenever a price level is updated, a history entry containing information about its previous state is created. It also includes information on the time period over which the state was effective.

In line with Xetra EnBS, we consider each delta message LOB update instruction to bring the LOB in a new state. In this context, we define two types of state. First, we consider the book to be in a final state when all updates reported in the last delta message have been integrated. In such a final state $\omega_i \in \Omega$, we expect the LOB to satisfy the consistency conditions (1.9a) to (1.10c). Also, a liquid stock order book being generally deeper than its twenty visible price levels for each side, our $L_i(ask)$ and $L_i(bid)$ functions should return twenty. On the other hand, we define $\omega_j \in \Omega$ as an intermediate state if at least one update reported in the last delta message has been integrated while at least one other still have to be. In such a state, depending on the instructions already assimilated, it is possible for the visible part of an order book side $s \in BS$ to contain a smaller number of price levels, which is tracked by $L_j(s)$. Additionally, we cannot assume the LOB to satisfy the consistency conditions in the intermediate states. But, since we work on a price level basis, the complete LOB consistency is not essential in order to implement our methodologies.

On a similar topic, the limited number of price levels included in the visibility window involves liquidity existing outside its limits. We have to make the distinction between this *not visible liquidity* that is related to a technical limitation and the *hidden liquidity* related to hidden and iceberg orders. It is possible for the former liquidity type to become visible if, for example, price levels present in the visibility window are cancelled or executed. On the other hand, despite the fact that it is sometime detectable via some trades execution analysis, no information about hidden liquidity is made available by Xetra EnBS. Consequently, given a price $p \in \mathbb{R}^+$ for which both visible and hidden liquidity is available on state $\omega_i \in \Omega$, it is an actual fact that $noo_i(p)$ and $qty_i(p)$ only provides information on the visible liquidity.

Before entering the core of our state tracking methodology, we complete our model by adding an actual time dimension. Therefore, we define the following function which relates a price to the timestamp of the last modification of its associated number of orders, aggregated quantity of shares or level number. This function establishes a link between our state based model and the real clock.

(1.11)	$\tau_i : p \in \mathbb{R}^+ \to \mathbb{R}^+$

All the required elements being available, the following expression defines the set of all possible price level states:

(1.12)	$STATE = \{(s, l, p, n, q, t_1, t_2) \in BS \times \mathbb{N}^+ \times \mathbb{R}^+ \times \mathbb{N}^+ \times \mathbb{N}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \mid l \leq 20\}$

For any price level state $(s, l, p, n, q, t_1, t_2) \in STATE$, s, l, p, q and n respectively correspond to the LOB side, level number, price, number of orders and aggregated quantity of shares characterizing the price level over the time period going from t_1 to t_2. We claim that a price level state is related to two LOB state changes. The first update occurs at t_1 and causes the price level to enter the state while the second, arriving at t_2, causes it to leave the state. Between these two events, the price level remains the same. The fact that a delta message can consist of several updates allows t_1 and t_2 to be identical. Since these states only exist for a Xetra EnBS technical reason, we discard them from our resulting dataset.

Order book changes being communicated through delta message update instructions, we synchronize our price level state tracking procedure with their arrival. Therefore, we define ω_i and $\omega_{i+1} \in \Omega$ as the states prevailing immediately before and immediately after a price level is updated. In this context, expressions (1.13) to (1.16) describe the relationship between states ω_i and ω_{i+1} elements in situations for which a price level is created, modified, or deleted. Although simple relations, these expressions can also be considered as a very general algorithm. For each of these LOB updates, we also present how to include our state saving mechanism and which elements are involved.

Based on definition (1.1), the following expressions present the effects of a new order book element $(t^n, s^n, l^n, p^n, n^n, q^n) \in NEW$ applied on the state ω_i LOB :

(1.13a)	$\forall k \in \mathbb{N}^+ \ s.t.\ k \geq l^n \ \wedge k \leq L_i(s^n),$ $n_{i+1}\big(p_i(s^n, k+1)\big) = n_i\big(p_i(s^n, k)\big),$ $q_{i+1}\big(p_i(s^n, k+1)\big) = q_i\big(p_i(s^n, k)\big),$ $\tau_{i+1}\big(p_i(s^n, k+1)\big) = t^n,$ $p_{i+1}(s^n, k+1) = p_i(s^n, k)$
(1.13b)	$\forall k \in \mathbb{N}^+ \ s.t.\ k < l^n,$ $n_{i+1}\big(p_i(s^n, k)\big) = n_i\big(p_i(s^n, k)\big),$ $q_{i+1}\big(p_i(s^n, k)\big) = q_i\big(p_i(s^n, k)\big),$ $\tau_{i+1}\big(p_i(s^n, k)\big) = \tau_i\big(p_i(s^n, k)\big),$ $p_{i+1}(s^n, k) = p_i(s^n, k)$
(1.13c)	$p_{i+1}(s^n, l) = p^n$
(1.13d)	$n_{i+1}(p) = n^n$
(1.13e)	$q_{i+1}(p) = q^n$
(1.13f)	$\tau_{i+1}(p) = t^n$
(1.13g)	$L_{i+1}(s^n) = min(L_i(s^n) + 1\ , 20)$

Expression (1.13a) presents the effect on the price levels whose number is larger than the l^n, which are pushed up by one level. Expression (1.13b) shows that new price level insertion has no effect on existing levels for which the number is smaller than l^n. Expressions (1.13c) to (1.13f) present how a new price level impacts functions (1.4), (1.6),(1.7) and (1.11) domains and codomains from state ω_{i+1}. Finally, expression (1.13g) simply shows that $L_{i+1}(s^n)$ must take the new level into account.

In the context of a new element, the only price levels for which the previous state must be saved are those being pulled up as presented in expression (1.13b). For each level number $k \in \mathbb{N}^+$ such that $k \geq l^n \wedge k \leq L_i(s^n)$, the following expression presents the price level state object figuratively created in order to be physically saved in a database :

(1.14)	$state_k = \left(s^n, k, p_i(s^n,k), q_i\left(p_i(s^n,k)\right), n_i\left(p_i(s^n,k)\right), \tau_i\left(p_i(s^n,k)\right), t^n\right)$

In a second scenario, using equation (1.2), we assume our price level to be transformed by an element change instruction $(t^c, s^c, l^c, n^c, q^c, v^c) \in CHANGE$. As before, the following expressions show the link between state ω_i and ω_{i+1} elements in the occurrence of this type of order book update instruction :

(1.15a)	$n_{i+1}\left(p_i(s^c,l)\right) = n^c$
(1.15b)	$q_{i+1}\left(p_i(s^c,l)\right) = q^c$
(1.15c)	$\tau_{i+1}\left(p_i(s^c,l)\right) = t^c$
(1.15d)	$\forall k \in \mathbb{N}^+ \ s.t.\ k \neq l^c \wedge k \leq L_i(s^c)$ $n_{i+1}\left(p_i(s^c,k)\right) = n_i\left(p_i(s^c,k)\right),$ $q_{i+1}\left(p_i(s^c,k)\right) = q_i\left(p_i(s^c,k)\right),$ $\tau_{i+1}\left(p_i(s^c,k)\right) = \tau_i\left(p_i(s^c,k)\right),$ $p_{i+1}(s^c,k) = p_i(s^c,k)$
(1.15e)	$v_{i+1}\left(p_i(s^c,l)\right) = v^c$

This time, the only affected price level is the one targeted by the update instruction. Thus, expressions (1.15a) to (1.15c) present how its properties change between our two states while expression (1.15d) shows that the other price levels remain identical. Since it is the only price level whose state must be saved, expression (1.14) is applied for $k = l^c$. By the end, (1.15e) keeps a track of the potential executed quantity of shares that may be related to the price level.

Finally, based on equation (1.3), the delete element instruction $(t^d, s^d, l^d, v^d) \in DELETE$ produces the following effect on our model elements:

(1.16a)	$n_{i+1}\left(p_i(s^d, l^d)\right) = 0$
(1.16b)	$q_{i+1}\left(p_i(s^d, l^d)\right) = 0$
(1.16c)	$\forall k \in \mathbb{N}^+ \; s.t. \; k \geq l^d \wedge k < L_i(s^d),$ $n_{i+1}\left(p_i(s^d, k)\right) = n_i\left(p_i(s^d, k+1)\right),$ $q_{i+1}\left(p_i(s^d, k)\right) = q_i\left(p_i(s^d, k+1)\right),$ $\tau_{i+1}\left(p_i(s^d, k)\right) = t^d,$ $p_{i+1}(s^d, k) = p_i(s^d, k+1)$
(1.16d)	$\forall k \in \mathbb{N}^+ \; s.t. \; k < l^d,$ $n_{i+1}\left(p_i(s^d, k)\right) = n_i\left(p_i(s^d, k)\right),$ $q_{i+1}\left(p_i(s^d, k)\right) = q_i\left(p_i(s^d, k)\right),$ $\tau_{i+1}\left(p_i(s^d, k)\right) = \tau_i\left(p_i(s^d, k)\right),$ $p_{i+1}(s^d, k) = p_i(s^d, k)$
(1.16e)	$L_{i+1}(s^d) = L_i(s^d) - 1$
(1.16f)	$v_{i+1}\left(p_i(s^d, l)\right) = v^d$

In this last case, a deleted price level can be interpreted as the removal of its available liquidity. Therefore, expressions (1.16a) and (1.16b) show that no liquidity remains for the price characterized by the provided book side and level number. Expression (1.16c) presents how price levels with numbers larger than the deleted price level are pushed down by one level. Expression (1.16d) shows that price levels for which the number is smaller than l^d do not change. Expression (1.16e) simply shows that the number of visible levels is decreased by one. Finally, (1.16f) keeps a track of the potentially executed volume reported concurrently for the price level.

This time, the price levels for which the state must be saved are the one being deleted and those being pushed down. In this context, we apply expression (1.14) on each $k \in \mathbb{N}^+$ such that $k \geq l^d \ \wedge k < L_i(s^d)$ in order to save the affected levels previous state.

The results of our states tracking methodology can be used in different ways. Let us first assume the set $SS \subseteq STATE$ corresponding to all saved price level states for a given trading day. As an example, it is possible to extract the following subset of SS which, once ordered using the t_1 value, corresponds to the time series of the best offer price level information: $\{(s, l, p, n, q, t_1, t_2) \in SS | s = ask \wedge l = 1\}$. To illustrate this example, Table 1.2 presents this actual time series using real data from the BMW title on February 1st, 2013 between 9:00:27 and 9:00:29 am.

It is also possible to obtain the complete state of the LOB visible part for any moment. Using the previous definitions, the following expression extracts a complete picture of the LOB for a given timestamp $t \in \mathbb{N}^+$: $\{(s, l, p, n, q, t_1, t_2) \in SS | t_1 \leq t \wedge t_2 > t\}$. Table 1.3 illustrates this example by presenting all the order book state entries for BMW on February 1st, 2013, at 9:45:00.000 am.

1.4 Summary

In this chapter, after having introduced the Xetra 2013 data that will be used as the backbone of this thesis in Section 1.1, we have performed two main tasks. First, in Section 1.2, we have formalized the different message types observed in the Xetra Enhanced Broadcast system output. We have mainly focussed on the trades and the order book update process regarding the provided price levels number of orders and quantity of shares updates. Having performed this task, we are almost able to get rid of the proprietary Xetra data structure. Indeed, from this point, all the following methodologies and analysis can be adapted to data sources presenting similar characteristics. This makes us in a very good position for the LOB events identification and limit order tracking tasks that will be performed in chapters 2 and 3.

Second, as a first actual contribution of this thesis, in Section 1.3, we have presented a methodology allowing to obtain the limit order book state for a given Xetra stock at any time during a trading day.[3] Although we only use these data to obtain summary statistics in this thesis, they have already been used as a primary data source in the context of significant academic work. Indeed, in addition to some thesis produced by students affiliated to the Canada Research Chair in Risk Management, they have been used as the main dataset for Dionne and Zhou (2020) and Dionne, Pacurar and Zhou (2015).

[3] Although it is outside of this thesis scope, we have been able to adapt this methodology to obtain similar information regarding the TSX securities for a period going from March 1, 2015 to August 31, 2015 (microseconds) and a period going from January 1 2016 to September 30, 2016 (nanoseconds).

Table 1.1 EnBS Xetra action examples

Initial LOB State					Limit order book event					Final LOB State			
Example 1 : New order book element													
Time	Level	Price	Qty	Orders	Action	Level	Price	Qty	Orders	Level	Price	Qty	Orders
9:00:30.684	1	74.07	40	1	New	3	74.13	509	1	1	74.07	40	1
	2	74.10	544	3						2	74.10	544	3
	3	74.14	441	1						**3**	**74.13**	**509**	**1**
	4	74.16	600	1						4	74.14	441	1
	5	74.20	53	1						5	74.16	600	1
Example 2 : Order book element change													
Time	Level	Price	Qty	Orders	Action	Level	Price	Qty	Orders	Level	Price	Qty	Orders
9:00:30.903	1	74.07	40	1	Change	1	74.07	63	2	1	74.07	**63**	**2**
	2	74.10	544	3						2	74.10	544	3
	3	74.13	509	1						3	74.13	509	1
	4	74.14	441	1						4	74.14	441	1
	5	74.16	600	1						5	74.16	600	1
Example 3 : Order book element delete													
Time	Level	Price	Qty	Orders	Action	Level	Price	Qty	Orders	Level	Price	Qty	Orders
9:00:31.225	1	74.07	63	2	Delete	4	n/a	n/a	n/a	1	74.07	63	2
	2	74.10	544	3						2	74.10	544	3
	3	74.13	509	1						3	74.13	509	1
	4	74.14	441	1						**4**	74.16	600	1
	5	74.16	600	1						**5**	**74.20**	**53**	**1**

Limit order book action examples

Table 1.2 Historical state entries examples

Level	Price	Quantity	Orders	Start Time	End Time
3	74.37	1200	1	9:00:27.155	9:00:27.241
3	74.33	740	1	9:00:27.241	9:00:27.243
3	74.26	223	1	9:00:27.243	9:00:27.421
3	74.25	250	1	9:00:27.421	9:00:28.012
3	74.24	323	2	9:00:28.012	9:00:28.028
3	74.24	223	1	9:00:28.028	9:00:28.037
3	74.24	323	2	9:00:28.037	9:00:28.039
3	74.23	100	1	9:00:28.039	9:00:28.424
3	74.23	323	2	9:00:28.424	9:00:28.424
3	74.23	530	3	9:00:28.424	9:00:28.430

Historical state entries examples.

Table 1.3 Order book state examples

Side	Level	Price	Quantity	Orders	Start Time	End Time
Ask	1	74.88	195	2	9:44:46.191	9:45:00.292
Ask	2	74.89	456	4	9:44:58.212	9:45:01.089
Ask	3	74.90	859	8	9:44:58.212	9:45:01.001
Ask	4	74.91	920	9	9:44:46.300	9:45:15.460
Ask	5	74.92	783	6	9:44:46.302	9:45:19.039
Ask	6	74.93	797	7	9:44:37.001	9:45:15.463
Ask	7	74.94	1200	7	9:44:57.836	9:45:18.986
Ask	8	74.95	2433	10	9:44:57.836	9:45:18.986
Ask	9	74.96	737	7	9:44:57.840	9:45:18.989
Ask	10	74.97	1206	9	9:44:20.586	9:45:30.226
Ask	11	74.98	523	5	9:44:20.532	9:45:30.226
Ask	12	74.99	478	3	9:44:42.939	9:45:15.469
Ask	13	75.00	825	8	9:44:20.532	9:45:30.215
Ask	14	75.01	2150	1	9:44:20.532	9:45:30.226
Ask	15	75.02	359	3	9:44:20.532	9:45:30.226
Ask	16	75.03	900	1	9:44:20.532	9:45:30.226
Ask	17	75.04	1913	3	9:44:46.301	9:45:15.458
Ask	18	75.06	100	1	9:44:20.532	9:45:30.226
Ask	19	75.09	650	1	9:44:20.532	9:45:30.226
Ask	20	75.10	1263	1	9:44:20.532	9:45:30.226
Bid	1	74.86	705	3	9:44:46.215	9:45:30.166
Bid	2	74.85	1230	3	9:44:57.001	9:45:09.001
Bid	3	74.84	351	4	9:44:44.680	9:45:13.001
Bid	4	74.83	936	7	9:44:45.026	9:45:18.986
Bid	5	74.82	905	9	9:44:52.119	9:45:30.212
Bid	6	74.81	2575	9	9:44:12.337	9:45:30.212
Bid	7	74.80	901	8	9:44:11.256	9:45:30.212
Bid	8	74.79	1417	9	9:44:19.473	9:45:30.212
Bid	9	74.78	593	6	9:44:11.245	9:45:30.212
Bid	10	74.77	1229	5	9:44:11.257	9:45:28.784
Bid	11	74.76	190	3	9:44:19.465	9:45:30.212
Bid	12	74.75	150	1	9:44:11.245	9:45:30.212
Bid	13	74.74	177	2	9:44:11.245	9:45:30.212
Bid	14	74.73	1949	3	9:44:12.337	9:45:30.212
Bid	15	74.72	320	2	9:44:11.245	9:45:30.212
Bid	16	74.71	1327	1	9:44:11.245	9:45:30.212
Bid	17	74.70	284	2	9:44:11.249	9:45:30.212
Bid	18	74.69	85	1	9:44:11.256	9:45:30.212
Bid	19	74.68	1656	1	9:44:11.256	9:45:30.212
Bid	20	74.66	1413	1	9:44:11.256	9:45:30.212

Complete limit order book state example for BMW on February 1st, 2013 at 9:45:00.000 am.

Chapter 2
Liquidity events identification

A limit order life cycle generally begins with its submission and ends with its cancellation or total execution. From our point of view, detecting and characterizing these events naturally represents the first step of a complete order monitoring process. However, as claimed before, providing information regarding these events occurrence does not appear to be part of the Xetra EnBS objectives. Indeed, instead of direct limit orders information, Xetra EnBS diffuses limit order book information under an incremental form. We claim this system to be designed with a strong latency minimization goal in mind, which actually translates into using a minimum amount of data to provide LOB information. For any price involving available liquidity, the only provided information is the number of visible orders and the total quantity of shares distributed among these orders. However since these values represent the available liquidity, their variations can be considered as proxies for liquidity inflows and outflows. Consequently, we expect them to be highly related to the orders events that we attempt to detect and characterize. Indeed, the submission of a limit order increases the LOB available liquidity level while its cancellation or execution decreases it. Already more oriented toward financial goals, in this chapter, we develop a backward procedure that uses liquidity variations in order to detect and characterize various types of event affecting the limit order book. After some definitions, we first present how to obtain liquidity variations directly from the Xetra EnBS LOB update instructions. Second, we analyze how liquidity flows interact to produce price level liquidity variations. Third, we define the set of liquidity events that we want to identify. We then develop a comprehensive set of rules for detecting and characterizing liquidity events from liquidity flows.

Since Xetra EnBS does not provide individual orders information, we do not expect to be able to make the distinction between the standard and some specialized order types such as *Stop Market*, *Stop Limit*, *Trailing Stop*, *One-cancels-other orders* and so on. Therefore, to reduce the overall complexity level, we define two general order categories that should encompass all Xetra more specific order types. First, we use the terms *passive order* for any order that becomes a part of the limit order book before execution or cancellation. It then refers to orders for which the execution is not sure. We assume this category to include simple limit orders as well as more complex objects such as the visible part of iceberg orders. Since a passive order joining the LOB is not always the result of an actual submission, for the sake of generality, we refer to an order affected by this type of actions as *added*. Similarly, in the absence of a transaction, we denote a passive order leaving the LOB as a *removed*. On the other hand, we keep referring to a limit order consumed in a trade context as *executed* since these terms induce no ambiguity. Finally, we identify the *aggressive order* as any order submitted for immediate execution. This category covers classical market orders and any type of marketable limit orders, which even includes Stop Orders for which the limit is reached[4]. We define the only event related to this category as *submitted aggressive order*. The best we can do in order to characterize these events is to group the reported trades by timestamp and affected book side. Then, for each group affecting the bid (ask) side on a given time, we identify an aggressive order submitted in order to sell (buy) a number of shares corresponding to the group aggregated number of shares.

2.1 Liquidity variations

When it comes to liquidity variation, regardless of their book side and rank, all price levels are treated the same. Consequently, we define our events monitoring methodology assuming a single price level involved in a single state change. We denote its book side, level number and price as s, l and p. We assume its liquidity variation to take place between states ω_i and ω_{i+1}. We assume ω_i to be a final LOB state.[5] Previously defined functions $n_{i+1}(p), n_i(p), q_{i+1}(p)$ and $q_i(p)$ remains the core of our next methodologies by providing our price level total number of orders and quantities of shares for both states. However, since the actual depth level price p remains constant and has no actual effect from this point, to lighten the notation, we replace these terms by n_i, n_{i+1}, q_i and q_{i+1}. The actual price level liquidity variations are obviously defined by $\Delta n = n_{i+1} - n_i$ and

[4] Marketable limit order refers to an ask (bid) limit order submitted at a price lower (higher) or equal to the best bid (ask) price, which cause it to be immediately, at least partially, executed.
[5] For ω_i to be actually final, the LOB update instruction reporting our price level state change would have to be reported first in the Xetra EnBS delta message, which is highly possible

$\Delta q = q_{i+1} - q_i$. Working on the basis that it initially contains q shares through n orders, this price level creation would translate into $n_i = 0$, $q_i = 0$, $n_{i+1} = n$ and $q_{i+1} = q$, which involves $\Delta n = n$ and $\Delta q = q$. Conversely, the same depth level deletion would imply $n_{i+1} = 0$ and $q_{i+1} = 0$, which results in $\Delta n = -n_i$ and $\Delta q = -q_i$. Finally, assuming a change on its constitution whose number of orders and total number of shares would become n and q, the same price level available liquidity variation would correspond to $\Delta n = n - n_i$ and $\Delta q = q - q_i$. Despite the fact that these definitions are based on the previously defined Xetra EnBS delta messages *new*, *delete* and *change* instructions, these concepts could easily extend to any other price level aggregated system.

2.2 Liquidity flows

Before working on liquidity events, we analyze the theoretical relation between liquidity variations and liquidity flows. With respect to liquidity inflows, we define n^a as the number of added visible passive orders involved in a price level state change. On the other side, n^r and n^e represent the number of removed and executed orders, which correspond to liquidity outflows. Representing the quantities of shares making up these orders using q^a, q^r and q^e, the following expressions show how liquidity flows interact to form a price level liquidity variation :

(2.1)	$\Delta n = n^a - n^r - n^e$
(2.2)	$\Delta q = q^a - q^r - q^e.$

Keeping in mind that it is possible for multiple passive orders to be simultaneously added and removed, we use the following conditions to formalize the fact that any added or removed passive order must contain at least one share :

(2.3a)	$n^a = 0 \Leftrightarrow q^a = 0, \qquad n^a > 0 \Rightarrow q^a \geq n^a$
(2.3b)	$n^r = 0 \Leftrightarrow q^r = 0, \qquad n^r > 0 \Rightarrow q^r \geq n^r$

Executed liquidity exhibits slightly different characteristics since it is possible for a passive order to be partially executed. In this case, the concerned order technically remains a part of the price level until the cancellation or execution of its last share. Consequently, it is not included in Δn. Because of this possibility, we claim n^e to report the number of passive orders being totally executed. The following condition ensures that any of these order involves at least one share :

(2.3c)	$q^e \geq n^e.$

In order to define consistent events identification rules, we link the previous expressions to our Xetra EnBS environment. While Δn and Δq are directly accessible, it is not the case for all (2.1) and (2.2) right members. Indeed, we do not individually observe orders added and removed related elements n^a, n^r, q^a and q^r. On the other hand, q^e is obtained through transactions information provided inside ATP and delta messages. Finally, n^e represents a special case since although not directly observed, it is related to q^e in many ways.

To simplify the analysis of liquidity flows related to added and removed orders, we first assume the absence of executed liquidity. In a perfect information granularity context, a positive liquidity variation could only be related to an added passive order while a negative variation would be attributable to an order being removed. However, despite its microsecond time precision, it is possible for Xetra EnBS to concurrently report liquidity flows of both types. Because of this eventuality, we claim that Δn and Δq provide the *net result* of liquidity flows. Indeed, assuming concurrently reported inflows and outflows, a positive Δn value involves a number of added orders larger than the number of removed order. Similarly, a positive Δq value means that the total number of added shares is larger than the number or removed shares, no matter how many orders they belong to. Naturally, the exact opposite applies to negative Δn and Δq values.

Two main difficulties are related to concurrently reported liquidity inflows and outflows. First, using Xetra EnBS data, it is globally impossible to determine, with an absolute certainty level, if we are in a situation where simultaneous passive orders addition and deletion are being reported or not. Second, even when such a situation is identified, it is sometime impossible to distangle the two types of order flows. Although not life threatening in the current events detection context, these issues may have consequences for our next methodology that uses identified events as starting points for complete passive order tracking purposes, as presented in Chapter 3. However, since to some extent there is nothing that can be done about it, we are resilient and attempt to use all available information to minimize the number of problematic situations and their consequences on our methodologies. In order to circumscribe identifiable and manageable situations, we divide net liquidity flows related to added and removed passive orders into three categories. First, we have the net inflows presumably related to added passive orders only. In these case, the observed liquidity variations meet the following criteria:

(2.4a)	$\Delta n \geq 1 \wedge \Delta q \geq \Delta n.$

While both price level number of orders and quantity of shares increase, Δq is large enough for any new limit order to consist in at least one share. The second category refers to net liquidity outflows presumably related to one or more orders being removed from the price level. This type of liquidity variations verify the following conditions:

(2.4b)	$\Delta n \leq -1 \wedge \Delta q \leq \Delta n.$

The last category essentially consists in cases for which conditions (2.4a) and (2.4b) are not verified. These net liquidity flows belong to concurrently reported added and removed passive orders situations. The more obvious cases are those where the price level number of orders and quantity of shares move in opposite directions. It also encompasses any situation in which, assuming their aggregated number of shares to differ, an equal number of added and removed orders is reported. Although sometime impossible to figure out, we present events detection rules specific to these situations later in this section.

By combining conditions (2.4a) and (2.4b) with Δn and Δq definitions, it is possible to establish that concurrently added and removed orders situations may hide among our first two categories if their underlying liquidity flows verify one of the following conditions:

(2.5a)	$n^r > 0 \wedge n^a - n^r \geq 1 \wedge q^a - q^r \geq n^a - n^r$
(2.5b)	$n^a > 0 \wedge n^r - n^a \leq -1 \wedge q^r - q^a \leq n^r - n^a$

In these cases, liquidity inflows and outflows combine in an offsetting way that simply cannot be distinguished from the presence of only one of these flow types. To illustrate this issue, Table 2.1 presents a simple example in which, starting from the same price level, two different liquidity flows scenarios produce the same net result. While scenario A consists in a simple added passive order, the more complex scenario B involves one removed and two added orders. Since they both produce the same final result, by looking at Δn and Δq values, it is impossible to identify the actual underlying scenario. However, although impossible to prove using our dataset, we assume these situations incidence to be reduced by the fact that at least three passive orders must be added or removed from the price level over a time period short enough for them to be reported by the same Xetra EnBS delta message. Indeed, for multiple liquidity flows verifying condition (2.5a) or (2.5b) to appear as a single group of added (removed) orders, at least two added (removed) orders and one removed (added) order must be reported by the same Xetra EnBS price level update, which should help keeping the number of out of control situations relatively low.

Up to this point, we have used $v_{i+1}: (p) \in \mathbb{R}^+ \to \mathbb{N}$ to represent the number of price p shares executed between LOB states ω_i and ω_{i+1}. Because of the still constant price p, we lighten the notation by substituting $v_{i+1}(p)$ by the simplest v_{i+1} expression. Beginning with the case presenting the highest degree of certainty, we consider the most common $v_{i+1} = 0$ as a sufficient condition to conclude in the absence of executed liquidity. This frequent case translates into $n^e = 0$ and $q^e = 0$. On the other hand, we assume $v_{i+1} > 0$ to signal executed liquidity, which implies Δn and Δq to be potentially related to this type of liquidity outflow. Before concluding in a direct relationship between v_{i+1} and q^e, we must consider that Xetra EnBS does not only provide information on the visible executed liquidity. Indeed, it is also possible for hidden executed liquidity to be reported through these values. We actually use the following expression to break our traded volume down into two components:

(2.6)	$v_{i+1} = q^e + q^h$

In this simple expression, q^e corresponds to the number of executed visible shares presented in (2.2) and q^h denotes the number of executed hidden shares for which the price is also p. Since Xetra EnBS does not make the distinction between these two types of liquidity when reporting traded volumes, we have to use the execution priority rules and the price level information to establish their respective values. These rules state that when standing on the same price level, visible liquidity must be completely consumed before hidden liquidity execution begins, no matter the orders arrival times. Consequently, as long as it does not exceed the number of visible shares standing on the price level before the transaction, we assume $v(t,p)$ to report visible executed shares only. The following expressions estimate the number of executed visible and hidden shares included in a potential traded volume:

(2.7)	$\hat{q}^e = \begin{cases} v_{i+1}, & v_{i+1} < q_i \\ q_i, & v_{i+1} \geq q_i \end{cases}$
(2.8)	$\hat{q}^h = \begin{cases} 0, & v_{i+1} < q_i \\ v_{i+1} - q_i, & v_{i+1} \geq q_i \end{cases}$

While $\hat{q}^e$ provides the number of price p visible shares whose execution is reported on time t, $\hat{q}^h$ does it for hidden shares. It is important to notice that these expressions do not consider liquidity flows that may affect the price level immediately before the transaction if they are reported through the same Xetra EnBS delta message. Although such rare undetectable cases lead to $\hat{q}^e$ and $\hat{q}^h$ biases, we consider their consequences as limited.

2.3 Liquidity events

Having highlighted several links between liquidity variations and liquidity flows, we are in a good position to define various event types that can be identified using these elements. In order to formalize the concept, the following expression describe the set of all possible events, which are represented using 6-tuples :

(2.9)	$EVENT = \{(t, s, l, p, \hat{n}, \hat{q}) \in \mathbb{R}^+ \times BS \times \mathbb{N} \times \mathbb{R}^+ \times \mathbb{N} \times \mathbb{N}\}$

For each event $(t, s, l, p, \hat{n}, \hat{q}) \in EVENT$, we define $t, s, l, p, \hat{n}$ and $\hat{q}$ as the reported timestamp, order book side, price level number, price, involved number of orders and involved quantity of shares. While $\hat{n}$ and $\hat{q}$ are computed using the price level liquidity variation and traded volume information, t is extracted from the delta message reporting the update. As before, s, l and p correspond to our subject price level information. These objects are useful for identified events manipulation purposes. However, to simplify the notation, we refer to our events using their $\hat{n}$ and $\hat{q}$ quantities. We use the accentuations to highlight the difference between these deduced values and their underlying counterparts presented in expressions (2.1) and (2.2).

We define *liquidity added* and *liquidity removed* events to characterize any type of passive order arrival and non execution related removal including but not limited to order submission, new iceberg order visible part, order cancellation and order expiration. We expect these events to match the liquidity flows when the price level liquidity change reports a single passive order submission or cancellation. Since it is possible for a liquidity flow to represent more than one added or removed passive order, we assume a liquidity added (removed) event to involve $\hat{q}^a$ ($\hat{q}^r$) shares distributed among $\hat{n}^a$ ($\hat{n}^r$) passive orders. Although first appearing possible to split $\hat{n}^a$ and $\hat{n}^r$ into multiple single order liquidity events, we avoid this avenue since we do not have the needed information to isolate the number of shares included in each order. To consistently match the underlying liquidity flows, we assume $\hat{n}^a > 0$ ($\hat{n}^r > 0$) to be verified and, since each affected passive order must consist in at least one share, $\hat{q}^a \leq \hat{n}^a$ ($\hat{q}^r \leq \hat{n}^r$) must also be true. These conditions recall the fact that no matter the number of shares represented, it is impossible for a passive order to be partially submitted or cancelled. We add a caveat since it is possible for a limit order to be partially executed while entering the LOB. In this situation, our definition states that the passive order would only correspond to the portion of the order becoming part of the LOB. Thus, its arrival would be characterized by a liquidity added event. Similarly, although possible for a partially executed passive order to remain on the book, it cannot be partially removed.

At this point, we have to express some concerns about limit orders modifications, which represent a real possibility on the Xetra stock market. From the Xetra specifications, there are two types of orders modifications. First, assuming that the price of an order is modified, in addition, or not, to the quantity of shares involved, the concerned order is removed from its initial price level while a new order is concurrently created on another price level, which corresponds to the order new physical price. In this situation, from our external point of view, we observe an order removed event on the initial order price level and an order added event on the new order price level. Consequently, we cannot make the distinction between such order modification and a situation where an order cancellation and an order submission would be concurrently reported, although unrelated. In this context, since we cannot make the distinction between the two scenarios, in any case, we identify the results as concurrent *liquidity removed* and a *liquidity added* events. In the second case where the orders quantity of shares is modified but not the price, we find ourselves in a situation where it is not always possible to determine what really happened. Indeed, assuming that this type of order modification is the only reported event at a given time, we observe a change in the concerned price level number of shares ($\Delta q \neq 0$) but not in the standing number of orders ($\Delta n = 0$). Once again, we cannot make the distinction between this type of order modification and concurrently reported order submissions and cancellations and once again, when possible, we will report it as separated *liquidity removed* and *liquidity added* events. We acknowledge that this situation is not ideal but we have to keep in mind that like us, the Xetra market participants were not able to make the distinction.

We define the *liquidity executed* event in a different way since it has no exact counterpart in the Xetra trading model. Indeed, with this event type, we attempt to document transactions effects from passive orders point of view. Therefore, liquidity executed events characterize the consequences of aggressive orders submission son the price level visible liquidity. We report these events in situations where, for a given price level, one or more passive orders are affected by one or more aggressive orders. Since we track these events on a price level basis, it is frequent for a single aggressive order to result in several of our liquidity executed events. We consider a liquidity executed event to affect $\hat{q}^e$ shares. In line with the previously described liquidity outflow related to executed passive orders, we define $\hat{n}^e$ as its number of totally consumed passive orders. As introduced before, liquidity executed is different from our liquidity added and removed events in that it can partially affect a passive order. This is related to the fact that this event type is the result of an aggressive order whose size, in shares, does not have to match any passive orders size. Therefore, while $\hat{q}^e > 0$ must always be verified since it must involve at least one share, it is possible for an event to partially affect an order, which leads to $\hat{n}^e = 0$. It may also report a single totally consumed passive order ($\hat{n}^e = 1$) or multiple affected orders ($\hat{n}^e \geq 1$). Assuming the event to totally consume at least one passive order, the number of consumed shares must be consistent with the usual required minimum of one share per passive order ($\hat{q}^e \geq \hat{n}^e$).

Our *hidden liquidity executed* event is similar to the liquidity executed event category since they both report trades effect from the consumed liquidity point of view. However, since the only known information about an executed hidden order reside in the traded quantity of shares, we can only characterize these events using this value. Indeed, there exists no situation in which we can determine the number of affected hidden passive orders that would be denoted $\hat{n}^h$. In this context, we assume a hidden liquidity executed event to simply affect $\hat{q}^h > 0$ shares. We consider these situations as the only occasion where Xetra EnBS provides information on hidden liquidity, which must actually be executed in order to be partially revealed.

Finally, we relate *liquidity moved in* and *liquidity moved out* events to passive orders entering and leaving the visible part of the LOB because of the limited number of visible price levels imposed by Xetra EnBS. Our *liquidity moved out* event is unique since the concerned levels are not directly affected by a LOB update instruction. As claimed before, any price level created inside this window causes another level with lower price priority to become invisible. This process allows the number of visible price levels to remain the same. In this context, once all instruction included in a delta message have been processed, we identify a liquidity moved out event for each price level excluded in this way. On the other hand, each time a price level is completely removed from the visible part of the LOB because of its complete execution or the cancellation of its last passive order, a new level becomes visible. In this situation, since a liquid asset LOB is generally deeper than the section made available by Xetra EnBS, we assume the new level to be an already existing one moving in from outside the LOB visibility window. Consequently, we consider this event as a *liquidity moved in* rather than the previously described liquidity added. In order to detect liquidity moves, we use the current number of visible price levels. Based on the Xetra EnBS rules of operation, we identify a new price level whose initial level number $l > L_i(s)$ as involved in a *liquidity moved in* event. It is possible for some new liquidity reported among these price levels to be identified as moved liquidity. However, we consider these potential misidentifications to have virtually no impact on our results. We also assume their occurrence frequency to be reduced by the fact that the concerned price levels are generally located in the second half of the visibility window, where the activity intensity is fundamentally low.

2.4 Events identification rules

Using the previously defined concepts, Table 2.2 presents our events identification rules, which are detailed in the next subsections. In a complementary way, Table 2.3 presents real world events identification examples used to illustrate each rule.

2.4.1 Added and removed liquidity

Rules 1.1 to 1.6 are related to liquidity added and liquidity removed events identified in the absence of a reported traded volume $v_{i+1} = 0$. We expect most of these cases to be related to simple passive orders submission and cancellation. Under rule 1.1, the emergence of a *new* price level is identified as a *liquidity added* event. We interpret this situation as a submitted passive order with a price for which no visible liquidity is already present in the LOB. Example 1.1a presents an actual case that we understand as the submission of a single ask side passive order made of 354 shares with a 38.455 euro price. This event results in the creation of a new visible price level inside the current bid-ask spread. Consequently, this new price becomes the best ask and the bid-ask spread is decreased. Regarding example 1.1b, Δn suggests the submission of two passive orders creating a new price level that becomes the best bid. Due to the lack of additional information, it is impossible to establish whether these two orders are submitted by the same or two different market participants. As claimed before, instead of characterizing a distinct event for each added order by attempting to arbitrarily separate Δq, we record a single liquidity added event for which $\hat{n}^a = 2$. It is interesting to note that the new liquidity identified under this rule is not affected by the possibility of simultaneously reported added and removed passive orders. Indeed, the number of events leading to the net liquidity inflow characterized by Δn and Δq has no effect on the perceptible final price level composition. On the other hand, under rule 1.2, new liquidity is added on an already existing price level. Since concurrent added and removed passive orders are possible, condition (2.4a) represents the heart of this rule. Example 1.2 illustrates this situation with a price level number of orders and quantity of shares increasing in a way that is consistent with the arrival of two passive orders cumulating 243 shares. We consider this liquidity inflow to form a single liquidity added event occurring on an existing price level.

On the opposite side, still assuming traded volume absence, we associate a deleted price level to the cancellation of all its passive orders. As presented in rule 1.3, we describe this situation as a *liquidity removed event* affecting the *entire* price level. When occurring on the first level of a LOB side, this type of event results in a wider bid-ask spread and a new best bid or best ask price. Although generally involving a price level consisting in a single passive order as shown in example 1.3a, it may affect a price level composed of multiple orders as presented in example 1.3b. In this second case, as for the *liquidity added event* involving multiple orders, we can discuss the possibility, or not, that the cancellations of orders submitted by two different operators are declared simultaneously by the trading system. Again, the lack of additional information leaves the question open. On another recurrent subject, Rule 1.3 represents a second set of situations in which concurrently reported added and removed orders have no consequence since, no matter the actual net liquidity outflow composition, the resulting price level is empty. On the other hand, Rule 1.4 presents the case where a price level decreasing number of orders and quantity of shares suggests one or more passive order to be cancelled. Therefore, although still existing, the post-event price level visible liquidity is lower. We associate this liquidity outflow to a *liquidity removed event partially* affecting the price level. The situation being more sensitive with respect to the possibility of concurrently reported liquidity flows, these events are detected using condition (2.4b). By way of illustration, Example 1.4 presents a case that we identify as the cancellation of a passive order containing 88 shares.

Rule 1.5 encompasses the only situations where it is possible to distangle concurrently reported added and removed passive orders. With the exception of some cases in which $\Delta n(p)$ and $\Delta q(p)$ are inconsistent with a single liquidity inflow or outflow, they involve a price level number of orders and quantity of shares moving in different directions. As explained before, expressions (2.1) and (2.2) demonstrate that in the absence of a traded volume, it is impossible to distinguish concurrently reported arriving and leaving liquidity for a price level. These are situations where, as explained before, it is possible that one order modification or more are involved. However, we define this rule in order to exploit an exception that we observe for cases where $n_i = 1$ or $n_{i+1} = 1$. In these situations, we assume the price level state ω_i liquidity to completely vanish in a *liquidity removed event*, being replaced by state ω_{i+1} liquidity in a *liquidity added event*. Example 1.5a presents a case in which the price level number of orders does not change while its quantity of shares decreases by 200. Although it may involve an order for which the quantity of shares is modified downwards while its price remains the same, as explained before, we identify this situation as a single passive order made of 400 shares being replaced by another order consisting in 200 shares. Example 1.5b presents a situation in which a single 353 shares passive order is replaced by two orders cumulating 271 shares. This case is detected on the basis of the simultaneous number of orders increase and quantity of shares decrease. Similarly, Example 1.5c presents a situation that we interpret as two passive orders totalling 207 shares being cancelled and one order composed of 500 shares being submitted. In these last two examples, although an order modification is possibly involved, it becomes clear that the order book is affected by multiple concurrently reported events.

Finally, Rule 1.6 presents the only cases that are, to the best of our knowledge, unsolvable using Xetra EnBS data. Facing these equations and inequations systems in which the number of unknow values surpasses the number of known values, we consider several added and removed passive orders combinations to be possible. To the extent that we attempt to produce results presenting a high degree of certainty, we cannot choose one of these combinations as the truth. Therefore, we simply mark these cases as unidentified. Examples 1.6a to 1.6c present interesting cases in which, moving in different directions, the price level number of orders and quantity of shares leave no real clues on the actual underlying liquidity flow.

2.4.2 Executed liquidity

Events identification Rules 2.1 to 2.9 characterize situations where LOB updates are accompanied by executed volumes. Each of these cases involving visible liquidity execution, these rules cover a wide range of situations in which up to three simultaneous liquidity events are identified. Using expressions (2.2) and (2.7), we define the following equations to characterize the involved quantities of visible shares:

(2.10)	$q_{i+1} = q_i - \hat{q}^e + \hat{q}^x$

where $\hat{q}^x$ represents the shares related to a net liquidity flow that may affect the price level visible liquidity concurrently to the executed volume. In terms of underlying liquidity flow, it is possible to establish that $\hat{q}^x = q^a - q^r$. As before, since we do not observe q^a and q^r, we retrieve their net values using $\hat{q}^x = \Delta q - \hat{q}^e$. Therefore, $\hat{q}^x > 0$ and $\hat{q}^x < 0$ represent net inflow and outflow while $\hat{q}^x = 0$ relates to the absence of concurrent liquidity flow.

Rules 2.1 and 2.2 describe trivial situations in which the only event affecting the price level appears to be the transaction. Under the first rule, the price level is totally consumed by the trade while, under the second, it is only partially affected. The traded volume matching the price level liquidity variation $\Delta q = -v_{i+1}$, it is easy to see that $\hat{q}^x = 0$. In both situations, we identify a *liquidity executed event* for which the affected quantity of shares corresponds to $\hat{q}^e$. Since they occur alone, the number of orders affected by these events correspond to $\hat{q}^e = v_{i+1} = -\Delta q$. Example 2.1 presents a case in which the price level contains a single passive order consisting of 170 shares that is totally consumed in a transaction. It results into a single *liquidity executed event entirely* consuming the price level. Example 2.2 relates to a situation in which only 262 of the 5192 available shares are consumed. This time, we identify a *liquidity executed event partially* affecting the price level. Additionally, the price level number of orders remaining the same, we assume that no passive order is totally consumed by the transaction.

Rules 2.3 to 2.5 represent almost identical cases for which equation (2.10) leads to $\hat{q}^x < 0$, suggesting one or more passive orders execution and cancellation to be concurrently reported. In each of these cases, we identify a *liquidity executed* event and a *liquidity removed event*. Under rules 2.3 and 2.4, these events combination leads to the complete price level deletion. The only difference between these two cases is that for 2.3, the fact that the price level initially contains a single passive order allows us to establish that the trade partially affect this order while its remainder is subsequently removed from the price level. Under Rule 2.4, we do not have enough information to determine the exact number of executed and removed orders. Finally, Rule 2.5 is identical to Rule 2.4 except for the fact that the events only partially affect the price level, which still exists after their occurrence has been reported. Example 2.3 presents a real case in which 200 shares are executed concurrently to the cancellation of 80 shares, leading to the identification of a *liquidity executed* and a *liquidity removed* events. Corresponding to Rule 2.3, this first example suggests the price level only passive order to have been affected by both *liquidity executed* and *liquidity removed* events. We suspect that the passive order was first partially consumed by an aggressive order, then cancelled. The time gap between these two events being short enough for them to be reported in the same Xetra EnBS delta message makes this case very interesting. We actually rely on chance to produce this type of results. Indeed, since the fill-or-kill condition does not apply to Xetra limit orders, we doubt this type of synchronicity to be the result of a market mechanism. Consequently, assuming the passive and aggressive orders to belong to different traders, we claim that the passive order owner may have launched its cancellation without being aware of its partial execution. Example 2.4 presents a similar situation in which the complexity is slightly increased by the fact that 4 passive orders are affected by the events combination. From equation (2.10), it is straightforward to conclude in a liquidity executed event affecting 802 shares and a liquidity removed event affecting the other 104 shares. However, because of the aggregated nature of Xetra EnBS price level information, it is impossible to make the distinction between completely executed and cancelled passive orders. Finally, Example 2.5 presents a Rule 2.5 real case in which the execution of 285 shares is reported concurrently to the withdrawal of 460 shares. Once these two actions accounted for, a single passive order containing 40 shares remains on the price level. The number of orders decreasing by 1 suggests two possible explanations. First, in a scenario similar to Example 2.3, it is possible for a single passive order to be affected by both *liquidity executed* and *liquidity removed* events. This order would be first partially consumed by the submitted aggressive order, then, its remaining would be immediately cancelled. In the second scenario, the two events would affect two different orders. Although impossible to establish the sequence in which both events affect the price level with certainty, we know that the partially executed order must be the one with the highest execution priority at the time of the transaction. It is possible that this priority has been changed by the other order cancellation if it occurred before.

Situations meeting Rules 2.5 and 2.6 criteria lead to the identification of simultaneous *liquidity executed* and *liquidity added* events. Also well represented by equation (2.10), they represent the case for which $\hat{q}^x > 0$. Because of this liquidity inflow, both situations result in a nonempty price level. Under Rule 2.5, we assume the price level available liquidity to be totally consumed first $(q_i = v_{i+1})$, then refilled with one or more new passive orders. In this context, we attribute n_i totally executed passive orders to the *liquidity executed event* and n_{i+1} added orders the *liquidity added event*. On the other hand, under Rule 2.6, the available liquidity being only partially executed $q_i > v_{i+1}$, we cannot distinguish the number of orders involved in both events.

2.4.3 Hidden or iceberg orders potential involvement

In a first scenario, passive and aggressive orders submitted at times close enough to be reported by the same delta message may represent a possible explanation for the concurrent *liquidity executed* and *liquidity added* events identified under Rules 2.5 and 2.6. However, these cases may also be, in a second scenario, related to the presence of an iceberg order on the price level for which the peak volume, or its remaining, is first entirely consumed by the trade. As we know, iceberg orders consist of two parts : a peak volume, which is visible, and a hidden volume. Each time the peak volume is totally executed, it is replaced by another one that is extracted from the hidden volume, which remaining quantity is decreased by the size of the new peak volume. This process goes on until the hidden volume is totally consumed. Since it is characterized by its own size and execution priority, an iceberg order peak volume behave exactly like a simple *passive* *order*. For price and time priority purposes, the timestamp attributed to the first peak volume corresponds to the time of the complete iceberg order submission. Subsequent peak volumes are marked with priority timestamps that ultimately correspond to the time of their immergence from the hidden volume, which also match the previous peak volume total execution time. We use Examples 2.6 and 2.7 to illustrate how, iceberg orders may be involved in situations described by Rules 2.6 and 2.7. The first example presents a real case in which a single passive order consisting of 421 shares appears concurrently executed and replaced by another order made of 1000 shares. Assuming the presence of an iceberg order, the passive order initially consumed in the *liquidity executed event* would represent the remaining of the previous peak volume. And the replacing 1000 shares order would correspond to the iceberg order new peak. Similarly, Example 2.7 exhibits a case in which the LOB update leads to the identification of a 904 shares *liquidity executed event* and a 500 shares *liquidity added* event. Still assuming the presence of an iceberg order to explain the situation, the 500 added shares would represent this iceberg new visible peak while the previous peak, or its remaining, would have been totally included in the 1318 shares consumed by the trade. To complete the interpretation, we would relate the one unit price level number of orders decrease to the total consumption of another passive order, not related to the iceberg. From our point of view, it is impossible to choose between the concurrently submitted passive and aggressive orders over a very short time period and the iceberg order explanations with a perfect confidence level. However, when an iceberg order is involved, the market trading system mechanically ensures both events to be reported with the same microsecond timestamp, which could improve this scenario likelihood.

In situations identified under Rule 2.8, we detect the involvement of three liquidity flows that we characterize using the same number of events. First, the traded volume being large enough to entirely consume the initial visible liquidity, we catalogue a *liquidity executed event totally* affecting the price level. We characterize this event as involving q_i shares distributed among n_i totally executed passive orders. Second, from expression (2.8), we know that $q_i < v_{i+1}$ involves the execution of $\hat{q}^h = v_{i+1} - q_i$ price p hidden shares. We relate this action to a transaction aggressive order counterpart probably requiring more liquidity than what was initially standing on the price level visible part. These hypotheses are supported by Xetra specifications which state that hidden liquidity must be consumed before visible liquidity standing higher in the price priority structure. In this context, we identify a *hidden liquidity executed event* representing the extra $\hat{q}^h$ shares. These shares being part of one or more hidden orders, we do not have sufficient information to establish the number of totally executed passive orders related to this event. Finally, in addition to these two execution related events, we detect a *liquidity added event* whose effect is to restock the price level visible liquidity. Following the same logic, we assume this event to involve q_{i+1} shares distributed among n_{i+1} passive orders. As before, although possible for this last event to be unrelated to the liquidity execution, this chain of events can be well explained by the presence of iceberg orders. Example 2.8 clearly illustrate this possible outcome. The initial price level consisting of a single 200 shares passive order, in the absence of hidden liquidity, the 220 shares trade would have totally consumed this liquidity and 20 shares on at least one deeper level. However, the fact that these extra 20 shares have been executed at the same price highly suggests the presence of hidden liquidity. Moreover, the 180 shares passive order concurrently added suggests this hidden liquidity to belong to an iceberg order. In such a case, the initial 200 shares passive order present on the price level would represent this iceberg order peak volume, already partially executed or not. Therefore, the total consumption of this volume would normally lead to the immergence of a new peak volume. However, based on Xetra execution rules, it would be normal for the extra 20 shares required by the transaction to be directly punctured from this next peak. In this context, it would be consistent for the new 180 shares passive order to actually be this new peak volume, from which 20 shares would have already been executed. In this scenario, the iceberg order default peak volume size would be 200 shares.

Except for the absence of liquidity added event, situations meeting the Rule 2.9 criteria are in every way similar to those identified under Rule 2.8. The entire price level is consumed in a liquidity executed event involving q_i shares distributed among n_i passive orders. And the exceeding traded volume quantity is attributed to a hidden liquidity executed event involving $\hat{q}^h = v_{i+1} - q_i$ shares. Although possible, it is more difficult to conclude in an iceberg order participation in the chain of events because of the absence of concurrent liquidity inflow. For an iceberg order to be involved in the transaction, its current peak volume would have to be part of the liquidity executed event affecting the entire price level visible liquidity. And, if still existing, its hidden volume would have to be totally consumed in the hidden liquidity executed event, which would explain the absence of a new peak volume. However, it is also possible and maybe more probable for the trade volume part exceeding the price level visible liquidity to consume a partial, one, or even multiple simple hidden limit orders. In this context, although impossible to get the exact explanation, we consider that once again, our identified liquidity events provide sufficient information to see the possible implications of such a situation.

2.4.4 Events without visual impact

Rules 3.1 to 3.4 apply to situations in which a reported traded volume is not accompanied by Xetra EnBS LOB update information. In these cases, we assume the price level number of orders and quantities of shares to remain the same. Therefore, we consider ω_i as the final state effective immediately before v_{i+1} is reported. We also assume that ω_{i+1} is a virtual state immediately following ω_i such that $n_{i+1} = n_i$ and $q_{i+1} = q_i$.

Rule 3.1 presents the only case in which we consider the traded volume as unrelated to an actual visible liquidity flow. Expressions (2.8) and (2.10) suggest that $n_i = 0$, $q_i = 0$, $n_{i+1} = 0$ and $q_{i+1} = 0$, which involves the executed volume to only consist in hidden liquidity. Therefore, we identify a *hidden liquidity executed event* affecting v_{i+1} shares. We have to keep in mind that it remains possible for a quantity of visible shares corresponding to the traded volume to join the price level immediately before the aggressive order submission. In this case, the aggressive order would execute against a visible passive order that would not exist long enough to be signaled to market participants. However, we consider the unique execution of hidden liquidity scenario as more likely.

In situations meeting Rules 3.2 to 3.4 criteria, the fact that the price level number of orders and quantity of shares do not change does not mean that its composition remains the same. It actually appears that in addition to the transaction, the price level is affected by a net liquidity inflow. Its invariable state suggests these elements to have a null effect on the available visible liquidity. In the spirit of (2.1) and (2.2), we claim that actual n^a, n^r, n^e, q^a, q^r and q^e values lead to $\Delta n = 0$ and $\Delta q = 0$. Since $n_i > 0$ and $q_i > 0$, equation (2.2) suggests the traded volume to execute against some of this visible liquidity. Therefore, we conclude in $n^a - n^r = n^e$ and $q^a - q^r = q^e$, which result in a null number of orders and quantity of shares variations.

From equation (2.10), we establish that Rules 3.2 and 3.3 represent Rules 2.6 and 2.7 special cases in which the *liquidity added event* offset the *liquidity executed event*. Both events affect a number of shares corresponding to v_{i+1}. Under 3.2, since the whole price level visible liquidity is consumed, it is possible to establish that the two events involve the same number of passive orders, which is characterized by n_i. On the other hand, because the price level is only partially consumed in situations meeting Rule 3.3 criteria, we do not have sufficient information to establish the number of passive orders involved in each event. As for Rules 2.6 and 2.7, we may reasonably consider iceberg orders involvement in situations identified under these two rules. Examples 3.2 and 3.3 illustrate this hypothesis with real cases. Beginning with Example 3.2, we assume the 500 shares single passive order initially composing the price level to be the peak volume of an iceberg order. We assume that in a first step, the reported transaction, which consists in 500 shares, has totally consumed this iceberg order visible part. Then, in a second step, a 500 shares volume has emerged from the iceberg order invisible part to become its new peak volume. In a similar way, Example 3.3 presents a case in which we assume the first passive order in the execution priority to consist in 500 shares. In this scenario, this order could also be the peak volume of an iceberg order. Since the transaction consumes exactly 500 shares, we assume this peak volume to be entirely executed in a *liquidity executed event* partially affecting the price level. We also assume that it is simultaneously replaced by a new peak through a *liquidity added event* involving 500 shares. As before, both examples may also be unrelated to iceberg orders. In such scenarios, the liquidity executed and added events would be triggered independently from each others. However, the facts that they involve the same numbers of passive orders and quantities of shares and they occur at times close enough not to cause Xetra EnBS LOB updates raise questions about these substitute scenarios likelihood.

Finally, we consider Rule 3.4 as a Rule 2.8 subcase in which the *liquidity executed* and the *liquidity added events* offset each other in a way that produce no visual effect on the visible liquidity. In Example 3.4, it is possible for the initial 180 shares single order standing on the price level to represent a partially consume iceberg order peak volume. Therefore, we assume an aggressive order attempting to consume 200 shares to totally execute this order and 20 shares from the next peak volume, which lead to the identification of a *hidden liquidity executed event*. Consequently, we also assume that this new partially affected peak volume enters the price level in a *liquidity added* event involving a single passive order consisting in 180 shares. Once again, these events may be unrelated but from our point of view, the timing factor and the quantities matches make the iceberg order scenario likely.

2.5 Results

In this section, we apply our events identification rules to the stocks composing the DAX, MDAX and SDAX indexes over the time period going from February 1 to April 30, 2013. While excluding the technological stocks, these indexes are composed of the Frankfurt Stock Exchange 30 largest market capitalizations, followed by the 50 most important medium capitalizations, and so on for 50 small capitalizations. Table 2.4 presents an inventory of the identified events regarding these three major indexes components. Creating a sort of *liquidity balance sheet*, we record the liquidity inflows and outflows observed through these events. In this context, liquidity inflows, which correspond to the *liquidity added* events, are attributed positive values in terms of number of involved orders and quantities of shares. On the other hand, liquidity outflows, which correspond to *liquidity removed* and *liquidity executed* events, are attributed negative values. For financial research purposes, we primarily focus on the events occurring during the daily continuous trading sessions.

Beginning with Panel A, we notice that 97.0%, 97.0% and 97.2% of the identified liquidity added events relates to *A.1 – Visible orders submissions*. Identified through Rule

1.1 and Rule 1.2, we consider them as the standard cases where visible liquidity is added to a new or an existing price level, most likely through simple orders submission. Table 2.4 Panel B presents similar results regarding the identified liquidity removed events. It shows that 96.9%, 97.0% and 97.3% of these events belongs to the *B.1 – Visible orders cancellation* sub-category. Identified through Rule 1.3 and Rule 1.4, we relate these events to standard situations where visible liquidity is removed from the price level, most likely through simple limit orders cancellation. We find the distribution of events among pairs of rules (1.1, 1.2) and (1.3, 1.4) interesting. Indeed, it is possible to observe that 9.1%, 46.4% and 81.1% of order submissions lead to the creation of a new price level (Panel A - Section A.1 - Rule 1.1). In a very similar way, we note that 6.8%, 45.4% and 81.3% of orders cancellations lead to an empty price level (Panel B – Section B.1 - Rule 1.3). We consider these numbers consistent with a general decrease in liquidity, in terms of number of orders standing on the price levels, as we move from the DAX to MDAX and SDAX components. Using the example of added liquidity, they show that while limit orders are generally submitted at a price for which orders are already existing for the DAX components, the almost opposite is observed regarding the SDAX stocks, which reflects much less filled order books. The MDAX components are found between with close to half of the new liquidity arriving on new price levels and slightly more than half on already existing ones.

Regarding the identification rules leading to more marginal proportions of liquidity added events, Table 2.4 Panel A Section A.2 presents the situations where liquidity added events are reported concurrently with visible liquidity executed events and, in some cases, with hidden liquidity executed events. Identified through six different rules, we consider that these situations have a high potential of being related to the automatic renewal of iceberg orders peak volume. In absolute terms, we observe 373 866, 167 722, 49 651 of these events over the period of interest. However, they only represent 0.18%, 0.23% and 0.16% of the liquidity added events identified for the DAX, MDAX and SDAX stocks, which remains negligible.

Other very marginal situations, Panel B Section B.2 shows that liquidity removed events identified concurrently with liquidity executed events represent only 0.06%, 0.03%, and less than 0.01% of all liquidity removed events. As claimed before, the exact sequence of actions leading to these situations may be nebulous. It is actually impossible to exclude that in several cases, the simultaneous report of the involved events by the Xetra EnBS system could be a simple coincidence, particularly in periods of high activity. This could explain their proportion decrease as we move from the highly liquid DAX stocks to the less liquid SDAX stocks, for which they are almost inexistent.

Regarding the situations that may be related to modified limit orders for which the price remains unchanged, as explained before, we face a case in which we are able to make the distinction between the removed and the added event (Rule 1.5) and a case where we consider this distinction impossible (Rule 1.6). Regarding the first case, Table 2.4 Panel A Section A.3 reports the liquidity added events identified and Panel B Section B.3 reports the liquidity removed events. It is possible to observe that the liquidity added events counterparts account for 0.21%, 1.38% and 2.28% of all liquidity added events identified. Similarly, the liquidity removed events identified in this context represents 0.22%, 1.46% and 2.32% of the observed liquidity removed events. Second, regarding the cases where we are not able to make the distinction among the involved events, the information is reported in Table 2.4 Panel A Section A.4 and Panel B Section B.4. It shows that these lost events represent 2.60%, 1.43% and 0.38% of all liquidity added events and 2.80%, 1.51% and 0.39% of all liquidity removed events reported during the continuous trading sessions. Since they essentially represent the only situations in which we are not able to obtain all the information of interest, we consider very fortunate that they correspond to such small proportions of the whole set of identified events.

When it comes to liquidity executed events, Table 2.4 Panel C Section C.1 shows that 91.3%, 91.3% and 85.7% of the cases are identified through either Rule 2.1 where the entire price level is consumed or through Rule 2.2 where it is only partially affected. Thus, as for liquidity added and removed events, these numbers show that a vast majority of liquidity executed events are identified in the most straightforward situations. Although they are reported concurrently with liquidity removed events, the same applies to the liquidity executed events presented in Section C.2, which corresponds to 1.44%, 0.76% and 0.07% of all liquidity executed events. When it comes to the liquidity executed events reported through section C.3 to C.5, it is important to note that they are identified in situations potentially involving iceberg or hidden orders. Therefore, it is possible for some of the visible liquidity whose execution is observed through these events to be part of iceberg orders peak volume. However, the available information do not allow us to validate this possibility. On the other hand, we are able to affirm that this liquidity was visible at the time of its execution. Consequently, by adding up the proportions reported in sections C.1 to C.5, we find that 97.7%, 97.7% and 95.3% of the identified liquidity executed events relate to visible liquidity for the DAX, MDAX and SDAX stocks. In counterbalance, Section C.6 shows that 2.34%, 2.29% and 4.66% of liquidity executed events involves liquidity that was not visible in the order book at the time of the execution. As explained before, we consider that these situations have a high potential to be related to the presence of either the invisible volume of iceberg orders or a hidden orders.

Finally, for results validation purposes, we find important to consider all the order book events taking place during a trading day, despite the fact that some of them may be related to Xetra market mechanics. In this context, Table 2.4 Panel D presents the net results regarding such residual miscellaneous events. First, in order to take it into account in the final balance, Section D.1 reports the net liquidity flows related to the previously described undistangled events (Rule 1.6), both in terms of number of involved orders and quantity of shares. Second, Section D.2 reports the events that we have identified as empty order book fulfillments. These events take place after an empty book period and are used to replenish it with all unexpired visible limit orders existing at the time. Depending on their expiration characteristics, these orders may be *Good-for-day*, *Good-till-date* or *Good-till-cancel*. They mostly occur after the opening, intraday and closing call auctions, as well as volatility halt periods. Although they take the appearance of added liquidity events, we cannot consider them as actual market participants actions such as passive orders submissions. They are therefore excluded from the final events result sets but we find them important for global liquidity flows validation purposes.

Section D.3 presents similar information regarding the liquidity crossing the limit of the 20 price levels window provided by Xetra. Once again, we cannot consider these so called events as actual orders submissions and cancellations and discard them from the final results. Finally, Section D.4 presents the net liquidity flows related to miscellaneous events observed during the Continuous trading, Opening call auction, Intraday call auction, Closing auctions, Volatility Halt and Post-trading sessions. Regarding the Intraday call auction, Closing auctions and Volatility Halt periods, these events are mostly related to the order book becoming empty which once again, take the shape of liquidity removed events but cannot be considered as orders cancellations. The remaining part of this category represents more obscure events presenting inconsistent characteristics and taking place over periods during which no activity should normally be observed in the order book. As an example, it happens that we observe liquidity inflows for which the quantity of shares is positive, but the number of orders is zero during the call auction periods, which makes no real sense from our point of view. We attribute this kind of activity to internal trading system messages and, it is obvious that it would be inconsistent to consider them as actual market participants actions.

In Table 2.4 Panel E, we consider the liquidity flows related to all the observed events, even those irrelevant for research purposes, and present the global net balances in terms of number of orders and quantities of shares. It is important to note that they only take visible liquidity into account, which lead to the exclusion of the information gathered in Panel C – Section C.6. In terms of involved number of orders, we observe positive balances of 314 823, 45 569, and 1 037 orders. We consider these marginal deviations as expected since there exist situations where it has been impossible to determine the actual number of visible orders involved in particular events. On the other hand, with regard to the number of involved shares, we find very interesting to observe perfect matches between the visible liquidity inflows and outflows for the DAX, MDAX and SDAX components over the three month period of interest. This performance allows us to have complete confidence in the results produced by our events identification methodology.

2.6 Summary

In this chapter, we have developed several rules allowing for the identification of liquidity added, liquidity removed, and liquidity executed events, which we associate to actual order submissions, cancellations, and executions. We have presented multiple real-life examples making it possible for anyone to perform the same task on similar data.

After having implemented this methodology on Xetra data regarding the DAX, MDAX and SDAX for the period going from February 1 to April 30, 2013, we have been able to establish that most of the events are identified through the most standard rules. In fact, we have shown that 97.0%, 97.0% and 97.2% of the liquidity added events, which can be related to limit orders submissions, are performed in the very usual context where liquidity is added to a new or an existing price level. We find very similar statistics regarding the 96.9%, 97.0% and 97.3% of liquidity removed event, which appear to take the shape of standard limit orders cancellation. Finally, we find that 91.3%, 91.3% and 85.7% of the liquidity executed events appear to concern usual visible liquidity.

Despite very interesting for future research purposes, we find that only 0.18%, 0.23% and 0.16% of liquidity added events may be attributed to iceberg orders new peak volumes. When it comes to hidden liquidity execution events, no matter the context, these proportions slightly increase to 2.34%, 2.29% and 4.66%.

Table 2.1 Concurrent liquidity flows example

$n_i(p)$	$n_{i+1}(p)$	$q_i(p)$	$q_{i+1}(p)$	Scenario	Passive order event	$\Delta n(p)$	$\Delta q(p)$
3	4	250	350	A	Added	1	100
					Net	1	100
				B	Removed	-1	-150
					Added	1	70
					Added	1	180
					Net	1	100

This table presents two examples of concurrent liquidity events leading to the same net results.

Table 2.2 Limit Order Book Events Identification Rules

LOB update conditions								Identified Events				
Rule	n_i	n_{i+1}	q_i	q_{i+1}	Δn	Δq	v_{i+1}	Id	Event Type	Level context	$\hat{n}$	$\hat{q}$
1.1	$= 0$	> 0	$= 0$	> 0				1.1.1	Liquidity added	New	Δn	Δq
1.2	> 0	> 0	> 0	> 0	> 0	$\geq \Delta n$	$= 0$	1.2.1	Liquidity added	Existing	Δn	Δq
1.3	> 0	$= 0$	> 0	$= 0$			$= 0$	1.3.1	Liquidity removed	Entire	$-\Delta n$	$-\Delta q$
1.4	> 0	> 0	> 0	> 0	< 0	$\leq \Delta n$	$= 0$	1.4.1	Liquidity removed	Partial	$-\Delta n$	$-\Delta q$
1.5	$= 1$	$= 1$	> 0	> 0	$= 0$		$= 0$	1.5.1	Liquidity removed	Entire	n_i	q_i
	or							1.5.2	Liquidity added	New	n_{i+1}	q_{i+1}
	$= 1$	> 1	> 0	> 0	> 0	$< \Delta n$						
	or											
	> 1	$= 1$	> 0	> 0	< 0	$> \Delta n$						
1.6	> 1	> 1	> 1	> 1	> 0	$< \Delta n$	$= 0$	1.6.1	Undefined	N/A	N/A	N/A
	or											
	> 1	> 1	> 1	> 1	< 0	$> \Delta n$						
	or											
	> 1	> 1	> 1	> 1	$= 0$							
2.1	> 0	$= 0$	$= v_{i+1}$	$= 0$		$= -v_{i+1}$	> 0	2.1.1	Liquidity executed	Entire	$-\Delta n$	$-\Delta q$
2.2	> 0	> 0	$> v_{i+1}$	> 0		$= -v_{i+1}$	> 0	2.2.1	Liquidity executed	Partial	$-\Delta n$	$-\Delta q$
2.3	$= 1$	$= 0$	$> v_{i+1}$	$= 0$			> 0	2.3.1	Liquidity executed	Partial	0	v_{i+1}
								2.3.2	Liquidity removed	Entire	1	$-(\Delta q + v_{i+1})$
2.4	> 0	$= 0$	$> v_{i+1}$	$= 0$			> 0	2.4.1	Liquidity executed	Partial	UN	v_{i+1}
								2.4.2	Liquidity removed	Entire	UN	$-(\Delta q + v_{i+1})$
2.5	> 0	> 0	$> v_{i+1}$	> 0		$< -v_{i+1}$	> 0	2.5.1	Liquidity executed	Partial	UN	v_{i+1}
								2.5.2	Liquidity removed	Partial	UN	$-(\Delta q + v_{i+1})$

LOB update conditions								Identified Events				
Rule	n_i	n_{i+1}	q_i	q_{i+1}	Δn	Δq	v_{i+1}	Id	Event Type	Level context	$\hat{n}$	$\hat{q}$
2.6	> 0	> 0	$= v_{i+1}$	> 0			> 0	2.6.1	Liquidity executed	Entire	n_i	q_i
								2.6.2	Liquidity added	New	n_{i+1}	q_{i+1}
2.7	> 0	> 0	$> v_{i+1}$	> 0		$> -v_{i+1}$	> 0	2.7.1	Liquidity executed	Partial	UN	v_{i+1}
								2.7.2	Liquidity added	Existing	UN	$\Delta q + v_{i+1}$
2.8	> 0	> 0	$< v_{i+1}$	> 0			> 0	2.8.1	Liquidity executed	Entire	n_i	q_i
								2.8.2	Hidden liquidity executed	UN	UN	$v_{i+1} - q_i$
								2.8.3	Liquidity added	New	n_{i+1}	q_{i+1}
2.9	> 0	$= 0$	$< v_{i+1}$	$= 0$			> 0	2.9.1	Liquidity executed	Entire	n_i	q_i
								2.9.2	Hidden liquidity executed	UN	UN	$v_{i+1} - q_i$
3.1	$= 0$		$= 0$		$= 0$	$= 0$	> 0	3.1.1	Hidden liquidity executed	UN	UN	v_{i+1}
3.2	> 0		$= v_{i+1}$		$= 0$	$= 0$	> 0	3.2.1	Liquidity executed	Entire	n_i	q_i
								3.2.2	Liquidity added	New	n_i	q_i
3.3	> 0		$> v_{i+1}$		$= 0$	$= 0$	> 0	3.3.1	Liquidity executed	Partial	UN	v_{i+1}
								3.3.2	Liquidity added	Existing	UN	v_{i+1}
3.4	> 0		$< v_{i+1}$		$= 0$	$= 0$	> 0	3.4.1	Liquidity executed	Entire	n_i	q_i
								3.4.2	Hidden liquidity executed	UN	UN	$v_{i+1} - q_i$
								3.4.3	Liquidity added	New	n_i	q_i

This table presents our events identification rules. For each rule to be enforced, conditions on n_i, n_{i+1}, q_i, q_{i+1}, Δn, Δq and v_{i+1} have to be verified. It is however important to note that the conditions regarding n_i, n_{i+1}, q_i and n_{i+1} do not ensure a valid limit order book. As an example, for any price level to be in a valid state, no matter the rule, the price level total number of shares has to be equal or larger than its number of orders $q_i \geq n_i$.

Table 2.3 Limit Order Book Events Identification Examples

	Context				Price level update information							Identified Events			
Case	t	s	l	p	n_i	n_{i+1}	q_i	q_{i+1}	Δn	Δq	v_{i+1}	Event Type	Level Context	$\hat{n}$	$\hat{q}$
1.1a	09:00:02.140	ask	1	38.455	0	1	0	354	1	354	0	Liquidity added	New	1	354
1.1b	09:00:16.026	bid	1	38.390	0	2	0	524	2	524	0	Liquidity added	New	2	524
1.2	09:01:48.213	bid	1	38.360	1	3	99	342	2	243	0	Liquidity added	Existing	2	243
1.3a	09:00:16.126	bid	4	38.365	1	0	200	0	-1	-200	0	Liquidity removed	Entire	1	200
1.3b	09:01:48.210	bid	3	38.350	2	0	400	0	-2	-400	0	Liquidity removed	Entire	2	400
1.4	09:01:48.221	bid	1	38.360	3	2	342	254	-1	-88	0	Liquidity removed	Partial	1	88
1.5a	09:01:57.645	bid	3	38.270	1	1	400	200	0	-200	0	Liquidity removed	Partial	1	400
												Liquidity added	Existing	1	200
1.5b	09:07:50.704	bid	1	38.495	1	2	353	271	1	-82	0	Liquidity removed	Partial	1	353
												Liquidity added	Existing	2	271
1.5c	15:50:33.619	ask	1	38.220	2	1	207	500	-1	293	0	Liquidity removed	Partial	2	207
												Liquidity added	Existing	1	500
1.6a	09:14:25.003	bid	1	38.600	2	4	651	632	2	-19	0	Liquidity removed	N/A	N/A	N/A
1.6b	09:07:23.951	bid	3	38.515	4	3	8518	8685	-1	167	0	Liquidity removed	N/A	N/A	N/A
1.6c	9:01:55.810	ask	17	38.520	3	3	302	900	0	598	0	Liquidity removed	N/A	N/A	N/A
2.1	09:00:10.092	ask	1	38.405	1	0	170	0	-1	-170	170	Liquidity executed	Entire	1	170
2.2	09:00:03.226	bid	1	38.400	2	2	5192	4930	0	-262	262	Liquidity executed	Partial	0	262
2.3	09:33:16.034	bid	1	38.405	1	0	280	0	-1	-280	200	Liquidity executed	Partial	0	200
												Liquidity removed	Entire	1	80
2.4	09:12:37.024	ask	1	38.615	4	0	906	0	-4	-906	802	Liquidity executed	Partial	N/A	802
												Liquidity removed	Entire	N/A	104

	Context				Price level update information							Identified Events			
Case	t	s	l	p	n_i	n_{i+1}	q_i	q_{i+1}	Δn	Δq	v_{i+1}	Event Type	Level Context	$\hat{n}$	$\hat{q}$
2.5	09:15:01.692	ask	1	38.620	2	1	785	40	-1	-745	285	Liquidity executed	Partial	N/A	285
												Liquidity removed	Partial	N/A	460
2.6	09:02:45.901	bid	1	38.330	1	1	421	1000	0	579	421	Liquidity executed	Entire	1	421
												Liquidity added	New	1	1000
2.7	09:05:35.319	ask	1	38.390	3	2	1318	914	-1	-404	904	Liquidity executed	Partial	N/A	904
												Liquidity added	Existing	N/A	500
2.8	09:01:42.135	bid	1	38.360	1	1	200	180	0	-20	220	Liquidity executed	Entire	1	200
												Hidden liquidity executed	N/A	N/A	20
												Liquidity added	New	1	180
2.9	09:03:51.548	bid	1	38.400	1	0	132	0	-1	-132	988	Liquidity executed	Entire	1	132
												Hidden liquidity executed	N/A	N/A	856
3.1	09:00:29.040	ask		38.400	0	0	0	0	0	0	313	Hidden liquidity executed	N/A	N/A	313
3.2	09:05:35.196	ask	1	38.390	1	1	500	500	0	0	500	Liquidity executed	Entire	1	500
												Liquidity added	New	1	500
3.3	09:51:42.495	ask	1	38.650	3	3	5594	5594	0	0	500	Liquidity executed	Partial	3	500
												Liquidity added	Existing	3	500
3.4	09:01:42.137	bid	1	38.360	1	1	180	180	0	0	200	Liquidity executed	Entire	1	180
												Hidden liquidity executed	N/A	N/A	20
												Liquidity added	New	1	180

This table present the main LOB events identification rules examples.

Table 2.4 Global DAX, MDAX and SDAX identified liquidity events

Panel A : *Liquidity added events*

Rule	Event	Level context	Count (M)	Orders (M)	Shares (M)	Count (M)	Orders (M)	Shares (M)	Count (M)	Orders (M)	Shares (M)	Events count %		
			DAX			MDAX			SDAX			DAX	MDAX	SDAX
A.1 - Visible orders submission														
1.1	1.1.1	New	18.8	19.4	10 642	33.5	33.9	15 166	25.4	25.5	21 019	9.1%	46.4%	81.1%
1.2	1.2.1	Exist	181.2	187.9	101 060	36.5	37.4	13 279	5.0	5.1	5 020	87.9%	50.5%	16.1%
		Total	*200.0*	*207.2*	*111 703*	*70.0*	*71.3*	*28 445*	*30.4*	*30.6*	*26 039*	*97.0%*	*97.0%*	*97.2%*
A.2 - Potential iceberg orders new peak volume														
2.6	2.6.2	New	0.146	0.152	177.6	0.069	0.074	30.0	0.013	0.014	13.2	0.07%	0.10%	0.04%
2.7	2.7.2	Exist	0.049	UN	73.7	0.015	UN	6.4	0.004	UN	3.6	0.02%	0.02%	0.01%
2.8	2.8.3	New	0.123	0.127	122.0	0.048	0.049	15.4	0.016	0.017	9.3	0.06%	0.07%	0.05%
3.2	3.2.2	New	0.047	0.047	36.7	0.029	0.030	10.0	0.014	0.014	7.8	0.02%	0.04%	0.04%
3.3	3.3.2	Exist	0.001	UN	1.2	0.000	UN	0.190	0.000	UN	0.220	0.00%	0.00%	0.00%
3.4	3.4.3	New	0.009	0.009	5.2	0.006	0.006	1.4	0.002	0.002	1.2	0.00%	0.01%	0.01%
		Total	*0.374*	*0.335*	*416.4*	*0.168*	*0.158*	*63.4*	*0.050*	*0.047*	*35.3*	*0.18%*	*0.23%*	*0.16%*
A.3 - Potential orders modification - Added orders counterparts														
1.5	1.5.2	New	0.424	0.436	178.0	0.999	1.0	308.8	0.716	0.716	433.8	0.21%	1.38%	2.28%
A.4 - Undistangles events - Potential orders modification - Added orders counterparts														
1.6	1.6.1	NA	5.4	UN	UN	1.0	UN	UN	0.119	UN	UN	2.60%	1.43%	0.38%
TOTAL (A)			**206.1**	**208.0**	**112 297**	**72.2**	**72.5**	**28 817**	**31.3**	**31.4**	**26 508**	**100%**	**100%**	**100%**

Panel B : *Liquidity removed events*

Rule	Event	Level context	Count (M)	Orders (M)	Shares (M)	Count (M)	Orders (M)	Shares (M)	Count (M)	Orders (M)	Shares (M)	Events count %		
			DAX			MDAX			SDAX			DAX	MDAX	SDAX
B.1 - Visible orders cancellation														
1.3	1.3.1	Entire	13.0	-13.5	-5 954	31.0	-31.5	-13 968	25.0	-25.1	-20 836	6.8%	45.4%	81.3%
1.4	1.4.1	Partial	173.0	-182.2	-97 137	35.3	-36.1	-13 095	4.9	-5.0	-4 751	90.1%	51.6%	16.0%
		Total	*186.0*	*-195.8*	*-103 091*	*66.3*	*-67.6*	*-27 063*	*30.0*	*-30.1*	*-25 588*	*96.9%*	*97.0%*	*97.3%*
B.2 - Visible orders cancellation - Concurrent to liquididty executed events														
2.3	2.3.2	Entire	0.010	-0.010	-3.4	0.003	-0.003	-0.684	0.000	0.000	-0.103	0.01%	0.00%	0.00%
2.4	2.4.2	Entire	0.063	UN	-26.6	0.013	UN	-3.1	0.000	UN	-0.057	0.03%	0.02%	0.00%
2.5	2.5.2	Partial	0.050	UN	-28.2	0.007	UN	-1.7	0.000	UN	-0.087	0.03%	0.01%	0.00%
		Total	*0.123*	*-0.010*	*-58.2*	*0.023*	*-0.003*	*-5.5*	*0.000*	*0.000*	*-0.247*	*0.06%*	*0.03%*	*0.00%*
B.3 - Potential orders modification - Removed orders counterparts														
1.5	1.5.1	Entire	0.424	-0.430	-171.1	0.999	-1.0	-319.7	0.716	-0.716	-322.7	0.22%	1.46%	2.32%
B.4 - Undistangles events - Potential orders modification - Removed orders counterparts														
1.6	1.6.1	NA	5.4	UN	UN	1.0	UN	UN	0.119	UN	UN	2.80%	1.51%	0.39%
TOTAL (B)			**191.9**	**-196.2**	**-103 321**	**68.3**	**-68.6**	**-27 388**	**30.8**	**-30.8**	**-25 911**	**100%**	**100%**	**100%**

Panel C: ***Liquidity executed events***

Rule	Event	Level context	Count (M)	Orders (M)	Shares (M)	Count (M)	Orders (M)	Shares (M)	Count (M)	Orders (M)	Shares (M)	Events count %		
			DAX			MDAX			SDAX			DAX	MDAX	SDAX
C.1 - Visible liquidity executed														
2.1	2.1.1	Entire	4.8	-8.0	-3 549	1.8	-2.4	-414.4	0.269	-0.331	-227.3	56.2%	56.8%	46.4%
2.2	2.2.1	Partial	3.0	-1.5	-2 567	1.1	-0.318	-211.9	0.228	-0.039	-149.4	35.1%	34.5%	39.4%
		Total	*7.8*	*-9.4*	*-6 116*	*2.8*	*-2.8*	*-626.2*	*0.497*	*-0.370*	*-376.7*	*91.3%*	*91.3%*	*85.7%*
C.2 - Visible liquidity executed - Concurrent to liquidity removed events														
2.3	2.3.1	Partial	0.010	0.000	-2.5	0.003	0.000	-0.459	0.000	0.000	-0.075	0.11%	0.11%	0.04%
2.4	2.4.1	Partial	0.063	UN	-55.7	0.013	UN	-4.1	0.000	UN	-0.106	0.74%	0.41%	0.02%
2.5	2.5.1	Partial	0.050	UN	-61.5	0.007	UN	-1.7	0.000	UN	-0.097	0.58%	0.24%	0.02%
		Total	*0.123*	*0.000*	*-119.7*	*0.023*	*0.000*	*-6.3*	*0.000*	*0.000*	*-0.277*	*1.44%*	*0.76%*	*0.07%*
C.3 - Visible liquidity executed - Concurrent to potential iceberg orders new peak volume														
2.6	2.6.1	Entire	0.146	-0.224	-154.8	0.069	-0.098	-23.6	0.013	-0.018	-11.9	1.70%	2.23%	2.29%
2.7	2.7.1	Partial	0.049	0.000	-118.4	0.015	0.000	-6.7	0.004	0.000	-5.2	0.57%	0.48%	0.61%
3.2	3.2.1	Entire	0.047	-0.047	-36.7	0.029	-0.030	-10.0	0.014	-0.014	-7.8	0.55%	0.94%	2.38%
3.3	3.3.1	Partial	0.001	UN	-1.2	0.000	UN	-0.190	0.000	UN	-0.220	0.01%	0.01%	0.03%
		Total	*0.242*	*-0.271*	*-311.2*	*0.114*	*-0.127*	*-40.5*	*0.031*	*-0.031*	*-25.1*	*2.83%*	*3.68%*	*5.30%*
C.4 - Visible liquidity executed - Concurrent to potential iceberg orders new peak volume and hidden liquidity executed														
2.8	2.8.1	Entire	0.123	-0.181	-141.3	0.048	-0.062	-15.6	0.016	-0.020	-12.5	1.45%	1.56%	2.84%
3.4	3.4.1	Entire	0.009	-0.009	-5.2	0.006	-0.006	-1.4	0.002	-0.002	-1.2	0.10%	0.18%	0.42%
		Total	*0.132*	*-0.190*	*-146.4*	*0.054*	*-0.068*	*-16.9*	*0.019*	*-0.023*	*-13.6*	*1.55%*	*1.74%*	*3.27%*
C.5 - Visible liquidity executed - Concurrent to hidden liquidity executed														
2.9	2.9.1	Entire	0.043	-0.077	-78.7	0.008	-0.013	-4.4	0.006	-0.008	-6.3	0.50%	0.27%	0.98%
C.6 - Hidden liquidity executed														
2.8	2.8.2	UN	0.123	UN	-163.8	0.048	UN	-17.4	0.016	UN	-16.6	1.45%	1.56%	2.84%
2.9	2.9.2	UN	0.043	UN	-81.2	0.008	UN	-8.2	0.006	UN	-16.4	0.50%	0.27%	0.98%
3.1	3.1.1	UN	0.025	UN	-9.4	0.009	UN	-2.2	0.002	UN	-0.811	0.29%	0.28%	0.41%
3.4	3.4.2	UN	0.009	UN	-7.4	0.006	UN	-1.8	0.002	UN	-2.2	0.10%	0.18%	0.42%
		Total	*0.200*	*0.000*	*-261.8*	*0.071*	*0.000*	*-29.8*	*0.027*	*0.000*	*-36.0*	*2.34%*	*2.29%*	*4.66%*
TOTAL[1]			**8.5**	**-10.0**	**-6 772**	**3.1**	**-3.0**	**-694.3**	**0.580**	**-0.432**	**-422.0**	**100%**	**100%**	**100%**

Panel D: *Miscellaneous events net results*

Rule	Event	Level context	Count (M)	Orders (M)	Shares (M)	Count (M)	Orders (M)	Shares (M)	Count (M)	Orders (M)	Shares (M)	Events count %		
			DAX			MDAX			SDAX			DAX	MDAX	SDAX
D.1 - Net Undistangles events														
1.6	1.6.1	NA	5.4	-0.004	-353.4	1.0	0.002	-77.5	0.119	0.000	-6.8	100%	100%	100%
D.2 - Empty book fullfillment														
1.1	1.1.1	New	0.152	0.320	838.1	0.248	0.312	220.0	0.248	0.318	416.6	100%	100%	100%
D.3 - Liquidity crossing price level 20														
From above level 20			16.3	27.3	48 711	19.7	24.2	25 656	8.7	11.8	16 950	48.7%	49.1%	49.8%
To above level 20			17.2	-28.6	-50 160	20.4	-25.0	-26 228	8.8	-11.9	-17 114	51.3%	50.9%	50.2%
		Net results	33.5	-1.3	-1 449	40.1	-0.841	-572.4	17.5	-0.103	-163.5	100%	100%	100%
D.4 - Net residual events														
Continous trading			0.000	0.000	0.000	0.000	0.000	-0.090	0.005	0.004	-0.837	0.00%	0.12%	1.22%
Open auction			0.000	0.000	0.000	0.004	0.000	0.004	0.020	0.000	0.814	0.03%	1.43%	5.37%
Intraday auction			0.077	-0.312	-421.8	0.151	-0.180	-118.3	0.188	-0.151	-180.6	48.9%	52.7%	49.7%
Close auction			0.073	-0.190	-717.6	0.125	-0.157	-182.9	0.129	-0.148	-211.0	46.4%	43.9%	34.0%
Volatility halt			0.007	-0.022	-100.4	0.005	-0.006	-3.4	0.036	-0.027	-30.3	4.60%	1.83%	9.60%
Post-trading			0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.055	0.00%	0.00%	0.02%
		Net results	0.158	-0.524	-1 240	0.286	-0.342	-304.7	0.378	-0.321	-421.9	100%	100%	100%

Panel E: *Global net events balance*

	Events Count	Orders	Shares	Events Count	Orders	Shares	Events Count	Orders	Shares
	DAX			MDAX			SDAX		
GLOBAL NET RESULTS	**434 813 828**	**314 823**	**0**	**183 209 277**	**45 569**	**0**	**80 695 021**	**1 037**	**0**

Chapter 3
Limit orders and liquidity tracking

In the previous chapter, we have developed a liquidity events detection and identification methodology by establishing an extensive set of rules essentially based on variations in the number of orders standing on the price levels and their aggregated quantity of shares. However, we have not taken the temporal dimension of liquidity into account. Until now, we have considered liquidity providing and consuming events as punctual and disconnected although in fact, they are related through the involved passive orders, which act as liquidity vehicles. From a LOB point of view, the submission of a passive order corresponds to a liquidity injection while its cancellation or execution relates to a liquidity withdrawal. At this point, it becomes relevant to ask if weather or not, it is possible to follow a given passive order from its submission to cancellation or execution across a price level aggregated limit order book. To address this question, the main purpose of this chapter is to develop a passive order tracking methodology that will be tested on the three main Xetra indexes stocks. This new methodology will be based on the previous chapter liquidity events identification rules. Since several Xetra EnBS information and event occurrence aggregation issues have been addressed through their development, these rules represent the ultimate passive orders tracking entry point.

As claimed before, Xetra EnBS diffused information is not centered on individual orders but on aggregated order book price levels. Consequently, the main part of our methodology consists in creating and maintaining an order list on a price level basis. We use visible liquidity events and their context to deduce information regarding limit orders entering and leaving the list. We interpret liquidity added event as signals of the arrival of one passive order or more on the concerned price level. In these cases, general orders information such as a buy or sell type, price, arrival price level number, quantity of shares and arrival time are deduced from the identified events. Concurrently, the new passive orders become part of the price level order list. On the other hand, we use liquidity removed events as signal of passive orders cancellation. On its occurrence, we gather termination information such as the cancellation price level number and cancellation time.

On the occurrence of these events, the concerned orders are saved and removed from their respective price levels list. Similarly, liquidity executed events signal the complete or total execution of one passive order or more. In such case, we use the price level list to identify the affected orders on an arrival time priority basis. Totally executed limit orders are updated and saved with an execution time and price level number then removed from the list. On the other hand, the remaining quantity of shares related to a partially executed limit order is updated and these ones remain part of the order list.

Submissions, cancellations, and execution having already been addressed in the previous chapter, most of the procedure lies in tracking these orders once they are part of their respective price level list. Our ultimate objective is to keep a track of passive orders from their submission to their cancellation or total execution. In a world where a price level would contain a single passive order, this task would be trivial. However, since several orders generally stand on a price level, some situations may become challenging and even impossible to resolve with our desired certainty level. Inserting passive orders to a given price level list is straightforward. For this operation, we simply assign the lowest execution priority to the last passive order entering the price level. When it comes to liquidity execution, we encounter two situations. First, it is possible for the whole price level to be executed. In such case, we simply consider each order present in the list as totally executed and save their termination information in consequence. In the second situation where the price level is partially executed, we use our execution priority structure to determine which orders are totally and partially executed. At this point, if our limit order list matches the actual underlying Xetra limit order list perfectly for the price level, we do not expect any problem. However, because of the Xetra LOB update diffusion system, as claimed before, it is possible for updates with zero net results to have introduced some mismatch in our list at some points in time. We attempt to detect such situation as early as possible by validating the pre-execution price level, post-execution price level and liquidity executed event numbers of orders and quantities of shares against each other. Finally, the problems which seem most prevalent arise from orders cancellation. In this case, three main situations are possible. The first one where a single order is present on the price level is very simple since we consider this order as cancelled. In the second situation, the price level list contains several passive orders each presenting a unique remaining quantity of shares. Assuming our list to match the actual Xetra underlying passive order list for the price level, it should contain an order for which the remaining quantity of shares would match that reported by the liquidity removed event. In such case, our task consist in saving cancellation information for this passive order and removing it from the list, which is also very simple. The actually problematic case comes from the third situation where more than one of the price level passive orders present a remaining quantity of shares that matches that of the liquidity removed event. In this case, we become unable to track the price level passive orders anymore. Indeed, since it represents the only distinctive information for a tracked passive order, each time a choice would have to be made between two passive orders presenting the same quantity of shares, we are forced to consider our methodology to be in a dead-end situation. In such a case, the bad choice could lead to a mismatch in the priority structure and produce an invalid price level passive order list. Thus, in situations leading to multiple order list possible configuration, we prefer to assume that we have lost track of the passive orders present on the level and attempt to reset the identification procedure.

This chapter is organized in the following way : In Section 3.1, we elaborate our complete order tracking methodology. In Section 3.2, we present the tracking results obtained by applying our methodology to the DAX, MDAX and SDAX indexes components. Finally, we conclude in Section 3.3.

3.1 Methodology

In addition to the previously described potentially problematic situations, our passive orders tracking methodology presents some limitations which are also related to the opaque nature of the Xetra diffusion system. Most of these issues regard actual limit orders characteristics. As an example, the quantity of shares is a characteristic that we cannot establish with a perfect level of certainty since possible for an aggressive order to be partially executed before its remaining becomes an actual limit order that is part of a price level. However, this issue having no potential hazardous impact except for information accuracy, we use the liquidity added event quantity of shares as the initial value. Similarly, the real validity timeframe of an order is actually unknown. Indeed, in the context of a liquidity removed event, we consider impossible to establish whether a passive order is manually cancelled by its owner or is automatically cancelled by the trading system because of its actual expiration. It represents a minor issue since both outcomes result in the limit order withdrawal from the book. Also, as for the liquidity events, we are unable to identify limit order modifications. Indeed, a price modification is considered as a cancelled limit order on the origin price level and a submitted limit order on the new price level. [6] Finally, a limit order quantity modification taking place on a price level comprising more than one order unfortunately leads to a dead-end limit orders tracking situation, which is not the case if the modified order is standing alone on the price level. As clamed before, in this last case, the modification is considered as a cancellation followed by a submission, which is simple in order tracking terms.

3.1.1 Meta-order definition

Before entering the core of our passive orders tracking methodology, we need to define some elements that will be used for its formalization and revisit others from Chapter 2. As we know from the previous section, we observe situations where the submission of more than one passive order are reported concurrently for a given price level. In order to efficiently manage such orders and to be able to integrate them into price level passive orders inventory, we elaborate the concept of *meta-order*, which acts as a container that may represent one passive order or more. In fact, each price level passive order list will be represented as a set of meta-orders. In the case where it represents a single order, a meta-order becomes perfectly in line with the actual passive order. In the other case where it represents more than one passive orders, since they reach the same price level at times close enough to be communicated as the same by the trading system, at first, we consider the general arrival characteristics of each of these orders to be identical. Moreover, these passive orders will share the same volume bank until they are all cancelled, totally executed, or, in the best-case scenario, they become part of a favorable situation where it becomes possible to distangle one or more of them. As long as they remain together, we consider multiple passive order meta-order members as a single order for priority

[6] A Xetra limit order price modification results in the assignment of a new priority timestamp, which is not the case when the quantity of shares is the only modified attribute.

management. Indeed, since the exact priority structure among the orders included in a meta-order being unknown, we consider them as an indissociable block. Therefore, other meta-orders can be below or above in the price level priority structure but they cannot be below a meta-order member and above another one.

Since the process is the same for any of them, we need a single price level to define our passive orders tracking methodology. Thus, we only consider liquidity events affecting the passive orders standing on this price level. Using p to denote its price, as before, n_i and q_i represent its standing number of orders and aggregated quantity of shares. As before, i refers to LOB state $\omega_i \in \Omega$ such that $\Omega = \{\omega_i : i \in \mathcal{I}\}$ and $\mathcal{I} = \{1,2,3,\dots\}$. To avoid notation overlapping, we use the ˜ accentuation to identify any object related to the previously introduced meta-order concept. We define $\tilde{O}_i^p$ as the set of all meta-orders that belong to our price p depth level on state ω_i. By definition, the atoms of this set correspond to the individual meta-order. Based on set theory, we translate these objects properties into functions. We define $\tilde{n}_i^p\colon \tilde{O}_i^p \to \mathbb{N}$, $\tilde{q}_i^p\colon \tilde{O}_i^p \to \mathbb{N}$ and $\tilde{t}_i^p\colon \tilde{O}_i^p \to \mathbb{R}$ to represent the number of orders, quantity of shares and arrival time related to a given meta-order. Extending the $\tilde{n}_i^p$ functions to the members of the $\tilde{O}_i^p$ power set leads to the $\tilde{N}_i^p\colon \mathcal{P}(\tilde{O}_i^p) \to \mathbb{N}$ function defined by the following expression :

(3.1)	$\forall \tilde{S} \in \mathcal{P}(\tilde{O}_i^p), \qquad \tilde{N}_i^p(\tilde{S}) = \begin{cases} \sum_{\tilde{s} \in \tilde{S}} \tilde{n}_i^p(\tilde{s}), & \lvert\tilde{S}\rvert > 0 \\ 0, & \tilde{S} = \emptyset. \end{cases}$

The number of orders of a meta-order set corresponds to the sum of all its components. By extension, $\tilde{N}_i^p(\tilde{O}_i^p)$, the number of orders of the price level complete meta-order set equals to n_i, which correspond to that of the price level. We apply the same properties to the quantities of shares through the $\tilde{Q}_i^p\colon \mathcal{P}(\tilde{O}_i^p) \to \mathbb{N}$ function as defined by the next expression, which similarly involves $\tilde{Q}_i^p(\tilde{O}_i^p) = q_i$.

(3.2)	$\forall \tilde{S} \in \mathcal{P}(\tilde{O}_i^p), \qquad \tilde{Q}_i^p(\tilde{S}) = \begin{cases} \sum_{\tilde{s} \in \tilde{S}} \tilde{q}_i^p(\tilde{s}), & \lvert\tilde{S}\rvert > 0 \\ 0, & \tilde{S} = \emptyset. \end{cases}$

As for events identification, we need a single order book state change to cover most orders tracking situations. As before, we assume this change to occur between consecutive LOB states ω_i and $\omega_{i+1} \in \Omega$. Therefore, i relates to the situation taking place just before the occurrence of one liquidity event having an effect on the passive orders standing on the price level or more, while $i+1$ relate to the situation taking place immediately after. In this context, we work with two sets of mathematical objects based on the previous definitions. Because our methodology applies to a single depth level and by extension, a single price, from this point, we lighten the notation by avoiding the p, which is constant. While $\tilde{O}_i$, $\tilde{n}_i: \tilde{O}_i \to \mathbb{N}$, $\tilde{q}_i: \tilde{O}_i \to \mathbb{N}$, $\tilde{t}_i: \tilde{O}_i \to \mathbb{R}$, $\tilde{N}_i: \mathcal{P}\big(\tilde{O}_i\big) \to \mathbb{N}$ and $\tilde{Q}_i: \mathcal{P}\big(\tilde{O}_i\big) \to \mathbb{N}$ represent the pre-events meta-orders characteristics, $\tilde{O}_{i+1}$, $\tilde{n}_{i+1}: \tilde{O}_{i+1} \to \mathbb{N}$, $\tilde{q}_{i+1}: \tilde{O}_{i+1} \to \mathbb{N}$, $\tilde{t}_{i+1}: \tilde{O}_{i+1} \to \mathbb{R}$, $\tilde{N}_{i+1}: \mathcal{P}\big(\tilde{O}_{i+1}\big) \to \mathbb{N}$ and $\tilde{Q}_{i+1}: \mathcal{P}\big(\tilde{O}_{i+1}\big) \to \mathbb{N}$ does it for the post-events properties.

We utilize the previously identified liquidity events as entry points in the management of the price level meta-order list and its components. Working with Chapter 2 Table 2.2 rules as general guidelines, we use the values obtained for liquidity added, removed, and executed events as the main inputs of our passive orders tracking methodology. As we go through these rules, we move from simple situations involving a single liquidity event to more complex situations in which up to three of them are identified. To avoid duplicating information already revealed in the previous section, we assume the identified liquidity events characteristics as already known. We rely on the previously defined $\hat{n}^a$, $\hat{q}^a$, $\hat{n}^r$, $\hat{q}^r$, $\hat{n}^e$ and $\hat{q}^e$ values to obtain liquidity added, liquidity removed and liquidity executed events affected number of orders and quantity of shares. Since we focus on visible liquidity, we do not take hidden liquidity executed events identified through Rules 2.8, 2.9, 2.1 and 3.4 into account.

3.1.2 Liquidity added

As described previously, Rule 1.1 relates to a liquidity added event leading to the creation of a new price level. Under Rule 1.2, the similar event affects an existing price level to which new liquidity is appended. In both cases, we initiate our tracking procedure for the passive orders involved in the added liquidity through the definition of a new meta-order $\tilde{o}^a$ which number of orders, quantity of shares and arrival time are related to previously defined concepts through the (3.3) expressions.

(3.3)	$\tilde{n}_{i+1}(\tilde{o}^a) = \hat{n}^a, \qquad \tilde{q}_{i+1}(\tilde{o}^a) = \hat{q}^a, \qquad \tilde{t}_{i+1}(\tilde{o}^a) = t$

In addition to these characteristics, we keep a track of complementary information such as the arrival market status, order book side, price, and level number.

For both rules, the newly created meta-order $\tilde{o}^a$ becomes part of the price level tracked meta-order list whose new composition is represented by (3.4). The sole difference between the two cases stands in the fact that under Rule 1.1, the pre-event meta-order list is empty $\tilde{O}_i = \emptyset$, which is not the case under Rule 1.2.

(3.4)	$\tilde{O}_{i+1} = \tilde{O}_i \cup \{\tilde{o}^a\}$

3.1.3 Liquidity removed

Moving forward in Chapter 2 Table 2.2, Rule 1.3 and Rule 1.4 present the cases of liquidity removed events appearing to take place alone. As claimed before, because of the trade absence, we relate them to the cancellation of one passive order or more. Rule 1.3 reflects the simplest situations since no passive order remains on the price level afterward. On the other hand, we interpret Rule 1.4 as the cancellation of one or multiple orders while one or multiple orders remain part of the price level. While trivial in the sole events identification context, this case becomes challenging in terms of passive orders tracking. Indeed, since the only known information regarding the cancelled liquidity is the total number of orders and quantity of shares removed from the price level, it may be impossible to determine the exact identity of the concerned order(s). It is easy to imagine a situation where it would be unfeasible to establish which one of two passive orders, presenting the same number of orders and quantity of shares, is actually affected by the cancellation. In such case, we consider that arbitrarily or randomly selecting an order could have undesirable effects that would be inconsistent with our objective of tracking passive orders with a maximum level of certainty. In addition to the risk of assigning the cancellation context information to the wrong order, this could introduce inconsistency in our price level execution priority structure. To avoid these potential damages, we identify these situations and take appropriate measures that will be described later in this section.

Before focussing on specific cases, we define some additional concepts related to liquidity that will be used not only when a liquidity removed event is reported alone, but also when it is concomitant to one or more event types. Starting form the fact that at least one meta-order should stand in the pre-event list, we consider that a meta-order or a set of meta-orders represents a *liquidity removed solution* if its characteristics perfectly match those of a liquidity removed event. Because of the fact that the same meta-order $\tilde{o} \in \tilde{O}_i$ may represent more than one passive order $(\tilde{n}_i(\tilde{o}) > 1)$, we define two types of liquidity removed solutions. In the case of a liquidity removed *exact* solution, the concerned meta-order or set of meta-orders aggregated number of orders and quantity of shares perfectly match those of the cancelled orders. With $\mathcal{P}^+(\tilde{O}_i)$ representing the non-empty elements of the state ω_i price level meta-orders power set, for a liquidity removed event involving n passive orders and q shares, (3.5) defines the function $\widetilde{ES}_i^r : \{(n,q) \in \mathbb{N}^2 | n > 0 \wedge q > 0\} \rightarrow \mathcal{P}^+(\tilde{O}_i)$, which provides the set of all exact solutions.

(3.5)	$\widetilde{ES}_i^r(n,q) = \{\tilde{S} \in \mathcal{P}^+(\tilde{O}_i) \mid \tilde{N}_i(\tilde{S}) = n \wedge \tilde{Q}_i(\tilde{S}) = q\}$

In the case of a liquidity removed *non-exact* solution, the concerned meta-order or group of meta-orders aggregated number of orders and quantity of shares are compatible with those of the cancelled liquidity, but the match is imperfect. These characteristics must obviously be larger than those of the removed liquidity and it must be possible for a part of the passive-order(s) represented by the concerned meta-order(s) to be withdrawn without leaving the price level in an inconsistent state. Thus, such meta-order(s) set must include at least one meta-order representing more than one passive order so that at least one of these orders may be considered as cancelled while at least one other is considered as remaining on the price level afterward. To illustrate the spirit of such solution, we take the example of a meta-order $\tilde{o}$ representing two concurrently submitted passive orders

($\tilde{n}_i(\tilde{o}) = 2$) and a total of two hundred shares ($\tilde{q}_i(\tilde{o}) = 200$) that would be the sole meta-order standing on the pre-event price level ($\tilde{O}_i = \{\tilde{o}\}$). Based on this information, it is possible to see that $\{\tilde{o}\}$ would correspond to a non-exact solution in the context of a liquidity removed event for which $\hat{n}^r = 1$ and $\hat{q}^r = 150$. The validity of this solution would involve that $\tilde{o}$ represents two actual passive orders : an order A consisting in 150 shares and an order B consisting in 50 shares. It would also involve that order A would be cancelled in the liquidity removed event context while order B would remain part of the price level. As a result, because it would afterward represent order B only, the features of meta-order $\tilde{o}$ would evolve into $\tilde{n}_{i+1}(\tilde{o}) = 1$ and $\tilde{n}_{i+1}(\tilde{o}) = 50$. It is interesting to note that this example also illustrates a situation where it becomes possible to make the distinction between two concurrently submitted limit orders on an ex post basis. Although intuitive in this example, a meta-order or a set of meta-orders representing a liquidity removed non-exact solution must meet some conditions in order to discard every possible case that would leave the post-event price level and its meta-order list in inconsistent states. In this context, expression (3.6) generalizes the concept of liquidity removed non-exact solution by defining the function $\widetilde{NES}_i^r : \{(n,q) \in \mathbb{N}^2 \middle| n > 0 \wedge q > 0\} \rightarrow \mathcal{P}^+(\tilde{O}_i)$, which actually provides the set of these solutions for a liquidity removed event involving n passive orders and a total of q shares.

(3.6)	$\widetilde{NES}_i^r(n,q) = \left\{ \tilde{S} \in \mathcal{P}^+(\tilde{O}_i) \middle\| \begin{matrix} \tilde{N}_i(\tilde{S}) > n \wedge \tilde{Q}_i(\tilde{S}) > q \wedge \\ \left(\exists \tilde{S}' \in \mathcal{P}^+(\{\tilde{o} \in \tilde{S} \mid \tilde{N}_i(\tilde{o}) > 1\})\right) \\ (\exists n' \in \mathbb{N}) \\ (\exists q' \in \mathbb{N}) \\ [(3.6.1) \wedge \ldots \wedge (3.6.7)] \end{matrix} \right\}$

where

(3.6.1)	$n' = n - \tilde{N}_i(\tilde{S} \setminus \tilde{S}')$
(3.6.2)	$q' = q - \tilde{Q}_i(\tilde{S} \setminus \tilde{S}')$
(3.6.3)	$n' \geq \lvert\tilde{S}'\rvert$
(3.6.4)	$q' \geq n'$.

(3.6.5)	$n' \leq \tilde{N}_i(\tilde{S}') - \lvert\tilde{S}'\rvert$
(3.6.6)	$q' \leq \tilde{Q}_i(\tilde{S}') - (\tilde{N}_i(\tilde{S}) - n)$

By formalizing the liquidity removed non-exact solution concept, (3.6) first shows that in order to represent this type of solution, the aggregated number of orders and the quantity of shares of the passive orders represented by the meta-order(s) $\tilde{S} \in \mathcal{P}^{+}(\tilde{O}_i)$ must be more important than those involved in the liquidity removed event.[7] It also shows that it must include a meta-order or group of meta-orders $\tilde{S}'$ representing more than one passive order. Then, (3.6.1) to (3.6.6) impose conditions in line with the fact that a valid non-exact solution have to include a part of each meta-order included in $\tilde{S}'$ while the other part is excluded. By controlling for the feasibility of such $\tilde{S}'$ meta-order(s) division, these propositions ensure that the application of a candidate solution would leave the price level in a consistent state. First considering the meta-order(s) included in $\tilde{S}$ but not in $\tilde{S}'$, if any, as totally included in the solution, (3.6.1) and (3.6.2) provide the $\tilde{S}'$ total number of orders and quantity of shares that would also be part of the solution while the residual part of each affected meta-order would remain in the list. Finally, expressions (3.6.3) to (3.6.6) ensure that the actual passive order(s) that would be included in the solution as well as those that would remain on the price level through $\tilde{S}'$ meta-order(s) fragmentation would exhibit valid properties, which is also necessary for the post-event price level consistency.

As a last step before linking the previous notions to our main methodology, we have to elaborate the procedure that will be invoked in any situation where it becomes impossible to track the price level passive orders. In such case, we first mark all the concerned state ω_i meta-orders as not trackable. Then, through the following expressions, we define a meta-order $\tilde{o}^m$ to regroup all passive orders standing on the state ω_{i+1} price level.

(3.7)	$\tilde{n}_{i+1}(\tilde{o}^m) = n_{i+1}, \qquad \tilde{q}_{i+1}(\tilde{o}^m) = q_{i+1}, \qquad \tilde{t}_{i+1}(\tilde{o}^m) = t\,.$

[7] Despite explicit in (3.6) for demonstrative reasons, from a purely logical point of view, these conditions are implicitly enforced through propositions (3.6.1) to (3.6.6).

This newly defined element becomes the only member of the meta-order list $\tilde{O}_{i+1} = \{\tilde{o}^m\}$. This mechanism enhances our results quality by allowing the execution priority structure to remain valid despite the fact that we have reached an order tracking dead-end situation.

Figure 3.1 uses the previously defined concepts to present how the liquidity removed events identified through Rule 1.3 and Rule 1.4 affect our passive orders tracking process. In both contexts, by applying the removed number of orders and quantity of shares $\hat{n}^r$ and $\hat{q}^r$ to (3.5) and (3.6), process (A) corresponds to the identification of any meta-order or group of meta-orders representing a valid exact or non-exact liquidity removed solution. Since no passive order remains on the price level after the occurrence of the event identified through Rule 1.3, this case involves the existence of a single liquidity removed exact solution since $\hat{n}^r = n_i$ and $\hat{q}^r = q_i$, which results into $\widetilde{ES}_i^r(\hat{n}^r, \hat{q}^r) = \{\tilde{O}_i\}$. Conditions (B) and (D) are consequently verified. This leads to the exact solution application (E), which marks a tracking process finality that we consider as a success. In the context of an exact solution application, for each meta-order present in $\tilde{S} \in \widetilde{ES}_i^r(\hat{n}^r, \hat{q}^r)$, we save the cancellation time, market status, price level number, as well as the executed and non-executed quantities of shares.[8] Afterward, since each passive order is considered as cancelled, the meta-order list new shape is provided by the next expression, which in the Rule 1.3 context, is equivalent to $\tilde{O}_{i+1} = \{\,\}$.

(3.8)	$\tilde{O}_{i+1} = \tilde{O}_i \setminus \tilde{S}$.

When it comes to Rule 1.4, Figure 3.1 (B) conditions become more solicited since in addition to the possibility of a single exact or a single non-exact solution, it is possible to face a dead-end situation where zero or more than one solution is consistent with the characteristics of the cancelled order(s). As presented by (C), in these two unfortunate cases, we apply the meta-orders merging procedure defined in (3.7). On the other hand, assuming the verification of the (B) conditions, the simplest case correspond to a situation

[8] It is important to note that in the case where an affected meta-order represents more than one passive order, these orders are saved together. Indeed, in such situation, we consider impossible to distangle these orders quantities of shares.

where a single exact solution $\left|\widetilde{ES}_i^r(\hat{n}^r, \hat{q}^r)\right| = 1$ exists and no non-exact solution $\left|\widetilde{NES}_i^r(\hat{n}^r, \hat{q}^r)\right| = 0$. Indeed, in such case, the identified exact solution is directly applied to the meta-order list through expression (3.8), which is represented by the process (E).

By contrast, a situation also meeting conditions (B) where a single non-exact solution is identified does not automatically lead to its application. Before applying a non-exact solution, we have to establish if it allows for a single cancelled liquidity allocation or not. Indeed, it is sometime possible for a group of meta-orders forming this type of solution to allow different ways to distribute $\hat{n}^r$ and $\hat{q}^r$ that result into a valid price level states. We identify these potential allocations as sub-solutions and the existence of more than one of them is incompatible with our order tracking objectives. However, since it appears impossible for a non-exact solution including a single meta-order representing more than one passive order to allow multiple sub-solutions, we use this criterion to determine the applicability of a non-exact solution, as expressed in (F). Therefore, given the existence of more than one sub-solution, the usual merging procedure (C) is applied. In the more desirable case where $\tilde{S} \in \widetilde{NES}_i^r(\hat{n}^r, \hat{q}^r)$ permits a single allocation and the meta-order $\tilde{o} \in \tilde{S}$ is the only meta-order representing multiple limit orders $\tilde{n}_i(\tilde{o}) > 1$, the solution takes effect in two stages. First, the meta-order $\tilde{o}$ evolution is reported by the following expressions :

(3.9)	$\tilde{n}_{i+1}(\tilde{o}) = \tilde{N}_i(\tilde{S}) - \hat{n}^r$ $\tilde{q}_{i+1}(\tilde{o}) = \tilde{Q}_i(\tilde{S}) - \hat{q}^r$.

Second, (3.10) provides the constitution of the price level meta-order list once the liquidity removed event applied :

(3.10)	$\tilde{O}_{i+1} = (\tilde{O}_i \backslash \tilde{S}) \cup \tilde{o}.$

3.1.4 Liquidity executed

As claimed before, the liquidity executed event is the second type of event potentially leading to the termination of a passive order tracking. We begin its coverage with the cases where it does not appear reported concurrently to any event of another type. Rule 2.1 represents the situation since the consumed liquidity exactly matches the price level visible liquidity and Rule 2.2 presents a situation where the price level visible liquidity is only partially executed. In line with the previous liquidity removed case, we introduce a concept of *liquidity executed solution* to identify the affected passive order(s) through the meta-order(s) representing them. However, in these cases, we have to take into account the execution priority structure and the fact that it is possible for a passive order to be partially executed.

Similar to the previously defined liquidity removed exact solution, we consider that a meta-order or a group of meta-orders represents a *liquidity executed exact solution* if its characteristics perfectly match those of a liquidity executed event. Such solution does not involve any order partial execution. We formalize this concept using the $\widetilde{ES}_i^e : \{(n,q) \in \mathbb{N}^2 | n \geq 0 \wedge q > 0\} \rightarrow \mathcal{P}^+(\tilde{O}_i)$ function, which is defined through the following expression and propositions :

(3.11)	$\widetilde{ES}_i^e(n,q) = \{\tilde{S} \in \mathcal{P}^+(\tilde{O}_i) \mid (3.11.1) \wedge (3.11.2) \wedge (3.11.3)\}$

where

(3.11.1)	$(\forall \tilde{o} \in \tilde{S})(\forall \tilde{o}' \in \tilde{O}_i \backslash \tilde{S})[\tilde{t}_i(\tilde{o}) < \tilde{t}_i(\tilde{o}')]$
(3.11.2)	$\tilde{N}_i(\tilde{S}) = n$
(3.11.3)	$\tilde{Q}_i(\tilde{S}) = q$

Using the arrival times, proposition (3.11.1) ensures that the passive orders represented by the solution meta-orders are first in line regarding our execution priority structure. Propositions (3.11.2) and (3.11.3) confirm the meta-order set to exactly match the liquidity executed event number of orders and quantity of shares.

We use the following function to determine the existence of a meta-order or a group of meta-orders forming a *non-exact liquidity removed solution*. Such solution involves the partial execution of the liquidity represented by one of the involved meta-order(s).

(3.12)	$\widetilde{NES}_i^e(n,q) = \{\tilde{S} \in \mathcal{P}^+(\tilde{O}_i) \mid P3.11.1 \wedge \ldots \wedge P3.11.6\}$

where

(3.12.1)	$(\forall \tilde{o} \in \tilde{S})(\forall \tilde{o}' \in \tilde{O}_i \backslash \tilde{S})[\tilde{t}_i(\tilde{o}) < \tilde{t}_i(\tilde{o}')]$
(3.12.2)	$\widetilde{N}_i(\tilde{S}) > n$
(3.12.3)	$\tilde{Q}_i(\tilde{S}) > q$
(3.12.4)	$(\forall \tilde{o} \in \tilde{S})[\tilde{Q}_i(\tilde{S} \backslash \tilde{o}) < q]$
(3.12.5)	$(\exists \tilde{o} \in \tilde{S})(\exists n' \in \mathbb{N})(\exists q' \in \mathbb{N})[\,(\forall \tilde{o}' \in \tilde{S} \backslash \tilde{o})[\tilde{t}_i(\tilde{o}) > \tilde{t}_i(\tilde{o}')]$ $\wedge\ n' = n - \widetilde{N}_i(\tilde{S} \backslash \tilde{o})$ $\wedge\ q' = q - \tilde{Q}_i(\tilde{S} \backslash \tilde{o})$ $\wedge\ n' \geq 0$ $\wedge\ n' \leq q'$ $\wedge\ \tilde{n}_i(\tilde{o}) - n' \geq 1$ $\wedge\ \tilde{n}_i(\tilde{o}) - n' \leq \tilde{q}_i(\tilde{o}) - q']$

Beginning with $\tilde{S} \in \mathcal{P}^+(\tilde{O}_i)$, the meta-order or group of meta-orders potentially representing a non-exact solution, relation (3.12.1) first ensures that this solution respects the execution priority structure. Second, (3.12.2) and (3.12.3) are used to verify that the number of orders and quantity of shares is large enough to cover the executed liquidity. Third, (3.12.4) ensures that in order to be valid, the solution requires each meta-orders contained in $\tilde{S}$. Finally, since a valid non-exact solution requires that a meta-order $\tilde{o} \in \tilde{S}$ is partially affected while the remaining meta-order(s) $\tilde{o}' \in \tilde{S} \backslash \tilde{o}$, if existing, are totally executed, relation (3.12.5) describes the conditions related to this partially affected meta-order $\tilde{o}$. The first (3.12.5) sub-relation ensures that, among all meta-orders involved in the solution $\tilde{S}$, $\tilde{o}$ is the last in line for execution. The second and third sub-relations establish the $\tilde{o}$ number of orders n' and quantity of shares q' considered as executed in the context of the solution. The fourth and fifth sub-proposition impose validity conditions on these quantities. In addition to ensuring the executed number of orders to be non-negative, they guarantee that each consumed passive order is linked at least to one consumed share. Finally, the last two sub-propositions ensure the validity of the meta-order $\tilde{o}$ non-executed liquidity. First they validate that at least one passive order is still represented by the meta-order. Then, they guarantee that each order contains at least one share. Through (3.12.5), we take into account that as long as a passive order is not totally consumed, it remains part of the price level number of orders and, in our context, of its representative meta-order. Indeed, a trade that only partially consumes one of the orders represented by $\tilde{o}$ must leave its parent meta-order number of orders unchanged.

Back to the two liquidity executed event cases described by Rule 2.1 and 2.2, it is important to note that unlike the previously described liquidity removed situations, because of the execution priority structure, each event should lead to the identification of a single solution. The fact that this solution is exact or non-exact depends on the match between the executed liquidity and the meta-order(s) characteristics. For the same reason, it is also impossible for a non-exact solution to include sub-solutions.

Very similar to the previously described Rule 1.3, Rule 2.1 situations directly lead to the identification of a single exact liquidity solution $\left|\widetilde{ES}_i^e(\hat{n}^e, \hat{q}^e)\right| = 1$ and consequently, the absence of non-exact solution $\left|\widetilde{NES}_i^e(\hat{n}^e, \hat{q}^e)\right| = 0$. The fact that the whole price level is executed involves that $\hat{n}^e = n_i$ and $\hat{q}^e = q_i$, which leads to $\widetilde{ES}_i^e(\hat{n}^e, \hat{q}^e) = \{\tilde{O}_i\}$. In this context, considering that all passive order have been executed, we close and save each meta-order present in $\tilde{S} \in \widetilde{ES}_i^e(\hat{n}^e, \hat{q}^e)$. The price level meta-order list evolves into (3.8), which in this specific case, is equivalent to $\tilde{O}_{i+1} = \{\,\}$.

When it comes to Rule 2.2, the fact that the price level is only partially executed leads to the possibility of an exact or a non-exact liquidity executed solution. Assuming an exact solution represented by $\tilde{S} \in \widetilde{ES}_i^e(\hat{n}^e, \hat{q}^e)$, we first save and close the involved meta-orders. Once again, the meta-order list content evolves into (3.8). On the other hand, assuming that $\tilde{S} \in \widetilde{NES}_i^e(\hat{n}^e, \hat{q}^e)$ corresponds to the liquidity executed non-exact solution and $\tilde{o} \in \tilde{S}$, which verifies $\left(\forall \tilde{o}' \in \tilde{S} \backslash \tilde{o}\right)[\tilde{t}_i(\tilde{o}) > \tilde{t}_i(\tilde{o}')]$, represents the partially affected meta-order, each meta-order potentially present in $\tilde{S} \setminus \tilde{o}$ are saved and closed. Moreover, in line with (3.9), while the number of orders characterizing $\tilde{o}$ evolves into $\tilde{n}_{i+1}(\tilde{o}) = \widetilde{N}_i(\tilde{S}) - \hat{n}^e$, its quantity of shares becomes $\tilde{q}_{i+1}(\tilde{o}) = \tilde{Q}_i(\tilde{S}) - \hat{q}^e$. The resulting price level meta-order list composition is described by (3.10).

Although unfortunate, the absence of both partial price level liquidity executed exact and non-exact solutions is made possible by an eventual mismatch between our reconstituted execution priority structure and the actual structure prevailing on the stock market. As claimed before, such discrepancies may be related to various out of control elements such as concurrent undistinguishable events reported for the same microsecond. Therefore, given the absence of liquidity executed solution $\left|\widetilde{ES}_i^e(\hat{n}^e, \hat{q}^e)\right| = 0$ and $\left|\widetilde{NES}_i^e(\hat{n}^e, \hat{q}^e)\right| = 0$, we are constrained to invoke our usual merging procedure, which creates a new starting point from which the execution priority structure validity is ensured.

3.1.5 More than one liquidity event

Having covered the rules related to the identification of a single type of event, we now focus on the situations where the occurrence of at least two types of events is reported simultaneously. Although more complex, in terms of order tracking, we generally handle each case by combining the previously defined concepts. We first move back to Rule 1.5 that relates to the succession of a liquidity removed event affecting the entire price level and a liquidity added event taking place on the new price level. Despite the fact that some of these situations are potentially related to simple passive order modifications with no effect on the price, the lack of information provided by Xetra forces us to process them as the cancellations and submissions of different orders. Consequently, the tracking process simply corresponds to the successive application of the Rule 1.3 and Rule 1.1 previously defined methodologies. When it comes to Rule 1.6, the unfortunate absence of consistent events information forces us to directly invoke our merging procedure.

Rules 2.3 to 2.5 describe situations leading to the identification of concurrently reported liquidity executed and a liquidity removed events. Because of the sole passive order standing on the pre-events price level, Rule 2.3 represents another simple case where the previously defined material directly apply. Indeed, it is possible to conclude in this order partial execution before the cancellation of its remaining. We manage this situation by combining Rule 2.2 and Rule 1.3 concepts. The Rule 2.2 methodology first lead to the identification of a non-exact liquidity executed solution involving $\hat{n}^e = 0$ passive order and $\hat{q}^e$ shares. Once applied, the methodology leads to the modification of the meta-order of interest from whose $\hat{q}^e$ shares are subtracted from it quantity of shares. Then, similar actions are performed using Rule 1.3 methodology that identifies an exact liquidity removed solution involving $\hat{n}^r = 1$ passive order and $\hat{q}^r$ shares. By applying this second solution, the concerned meta-order is considered as totally cancelled. It is consequently saved and closed, leading to an empty meta-order list for the price level.

When it comes to Rule 2.4 and Rule 2.5, our orders tracking methodology is complicated by the fact that we do not directly observe the number of complete order(s) consumed by the trade $\hat{n}^e$ and the number of removed order(s) $\hat{n}^r$. As seen in the previous sections, it is however possible to relate these values through the observed variation of the number of orders standing of the price level : $\Delta n = -(\hat{n}^e + \hat{n}^r)$. By combining this expression with some previously defined concepts, we extend our methodology and elaborate a procedure allowing, in some cases, to identify the passive order(s) involved in each of the two events. Before going further, we have to state on our inability to process situations where the pre-events meta-order(s) set consists in a single element representing more than one passive order $|\tilde{O}_i| = 1 \wedge \tilde{N}_i(\tilde{O}_i) > 1$. Such aggregation generally makes the determination of the exact events arrival sequence and its effects on the represented orders impossible. In these cases, we invoke our closing and merging procedure to preserve the validity of our tracking results. In any other situations we consider the existence of a single complete solution as desirable for passive orders tracking purposes. In this context, we verify if there exists zero, one or more valid transformation sequences which, starting from the pre-events meta-order set $\tilde{O}_i$, result in a post-event set $\tilde{O}_{i+1}$ that is consistent with the price level information. Since we cannot immediately determine which of the liquidity executed or the liquidity removed event have occurred first, we have to analyze both eventualities. We identify the case in which the liquidity removed event would have occurred before the liquidity executed event as *scenario A* and the reverse events sequence as *scenario B.*

As a first step in our attempt to establish if the event sequence corresponds to scenario A, we define expression (3.13) to obtain $\breve{N}_A^r$, a preliminary set of candidate values for the unobserved $\hat{n}^r$ quantity.

(3.13)	$\breve{N}_A^r = \left\{ \breve{n}_A^r \in \mathbb{N} \middle| \begin{array}{c} \breve{n}_A^r \geq 1 \ \wedge \\ \breve{n}_A^r \leq -\Delta n \ \wedge \\ \breve{n}_A^r \leq \hat{q}^r \ \wedge \\ \widetilde{ES}_i^r(\breve{n}_A^r, \hat{q}^r) \cup \widetilde{NES}_i^r(\breve{n}_A^r, \hat{q}^r) \neq \{\} \end{array} \right\}$

In addition to meet the removed order(s) validity conditions, when combined with the observed quantity of removed shares $\hat{q}^r$, each $\breve{n}_A^r \in \breve{N}_A^r$ value must lead to at least one liquidity removed solution, which is ensured using (3.5) and (3.6) functions.

In the case where $\left|\breve{N}_A^r\right| = 0$, we consider impossible to establish that the liquidity removed event have affected the price level before the liquidity executed event. Therefore, we simply move forward with the analysis of scenario B likelihood. On the other hand, the case where $\left|\breve{N}_A^r\right| \geq 1$ may lead to different outcomes in terms of order tracking. First, we have to consider that because it would encompass multiple sub-solutions, the identification of a potential non-exact solution involving more than one meta-order that represents multiple passive orders leads to an impasse. In this situation, we cannot go further with our order tracking process for the reasons previously enumerated. In formal terms, if $\exists \breve{n}_A^r \in \breve{N}_A^r$ such that $\exists \tilde{S}_A^r \in \widetilde{NES}_i^r(\breve{n}_A^r, \hat{q}^r)$ such that $\left|\left\{\tilde{o} \in \tilde{S}_A^r \middle| \tilde{n}_i(\tilde{o}) > 1\right\}\right| > 1$, we invoke our usual dead-end meta-orders merging procedure. In the more promising second situation where these criteria are not met, we have to determine if zero, one or more potential liquidity removed solution is consistent with an eventual subsequent liquidity executed solution. Since a candidate number of removed orders $\breve{n}_A^r \in \breve{N}_A^r$ may lead to multiple liquidity removed solutions $\tilde{S}_A^r \in \widetilde{ES}_i^r(\breve{n}_A^r, \hat{q}^r) \cup \widetilde{NES}_i^r(\breve{n}_A^r, \hat{q}^r)$, we have to analyze each of these eventual combinations individually. For each potential solution

$\tilde{S}_A^r$ identified for a candidate number of orders $\breve{n}_A^r$, we first create a parallel state $\omega_{i'}$ by virtually applying $\tilde{S}_A^r$ to the pre-events meta-orders set $\tilde{O}_i$. If the solution is exact $\left(\tilde{S}_A^r \in \widetilde{ES}_i^r(\breve{n}_A^r, \hat{q}^r\ )\right)$, this new set is simply defined as $\tilde{O}_{i'} = \tilde{O}_i \setminus \tilde{S}_A^r$. If the solution is non-exact $\left(\tilde{S}_A^r \in \widetilde{NES}_i^r(\breve{n}_A^r, \hat{q}^r\ )\right)$, assuming that $\tilde{o}_A^r \in \tilde{S}_A^r$ corresponds to the meta-order representing more than one order $(\tilde{n}_i(\tilde{o}_A^r) > 1)$, we define $\tilde{O}_{i'}$ as $\left(\tilde{O}_i \setminus \tilde{S}_A^r\right) \cup \tilde{o}_A^r$, $n_{i'}(\tilde{o}_A^r) = \tilde{N}_i\left(\tilde{S}_A^r\right) - \breve{n}_A^r$ and $q_{i'}(\tilde{o}_A^r) = \tilde{Q}_i\left(\tilde{S}_A^r\right) - \hat{q}^r$. Then, it becomes easy to establish the consistency of a liquidity executed solution involving $\hat{q}^e$ shares distributed among $\breve{n}_A^e = -\Delta n - \breve{n}_A^r$ passive orders by applying functions (3.11) and (3.12) on $\tilde{O}_{i'}$. By performing this sequence on each combination of a candidate number of removed orders and potential liquidity removed solution, it becomes possible to identify those leading to a valid liquidity executed solution. Unfortunately, if more than one of these successful combinations is identified, we cannot keep tracking the price level passive orders and invoke our dead-end merging procedure. Otherwise, we perform the next steps to analyze the feasibility of scenario B.

To determine the scenario B occurrence possibility, we use a methodology very similar to that developed for scenario A. Since the steps are simply reversed, we begin by using the following expression to obtain our preliminary set of candidate values for the unobserved number of totally executed orders $\hat{n}^e$.

(3.14)	$\breve{N}_B^e = \left\{ \breve{n}_B^e \in \mathbb{N} \,\middle\vert\, \begin{array}{c} \breve{n}_B^e \geq 0 \ \wedge \\ \breve{n}_B^e < \Delta n \ \wedge \\ \breve{n}_B^e \leq \hat{q}^e \ \wedge \\ \widetilde{ES}_i^e(\breve{n}_B^e, \hat{q}^e\ ) \cup \widetilde{NES}_i^e(\breve{n}_B^e, \hat{q}^e\ ) \neq \{\ \} \end{array} \right\}$

In the case where $\left|\breve{N}_B^e\right| \geq 1$, we perform our analysis by applying each combination of candidate number of totally executed orders $\breve{n}_B^e \in \breve{N}_B^e$ and potential liquidity executed solution $\tilde{S}_B^e \in \widetilde{ES}_i^e(\breve{n}_B^e, \hat{q}^e\ ) \cup \widetilde{NES}_i^e(\breve{n}_B^e, \hat{q}^e\ )$ to $\tilde{O}_i$, in order to characterize an alternative state $\omega_{i''}$. Thus, if $\tilde{S}_B^e \in \widetilde{ES}_i^e(\breve{n}_B^e, \hat{q}^e\ )$, the new meta-orders set $\tilde{O}_{i''}$ is defined as $\tilde{O}_i \setminus \tilde{S}_B^e$. Otherwise, if $\tilde{S}_B^e \in \widetilde{NES}_i^e(\breve{n}_B^e, \hat{q}^e\ )$, assuming $\tilde{o}_B^e$ to represent the non-exact solution partially executed meta-order, $\tilde{O}_{i''} = \left(\tilde{O}_i\ \setminus \tilde{S}_B^e\right) \cup \tilde{o}_B^e$, $\tilde{n}_{i''}(\tilde{o}_B^e) = \tilde{N}_i\left(\tilde{S}_B^e\right) - \breve{n}_B^e$ and

$\tilde{q}_{i''}(\tilde{o}_B^e) = \tilde{Q}_i(\tilde{S}_B^e) - \hat{q}^e$. Once the state i'' elements defined, from $\breve{n}_B^r = -\Delta n - \breve{n}_B^e$, we use the following expression to obtain a set of potential liquidity removed solutions.

$$\tilde{S}_B^r = \widetilde{ES}_{i''}^r(\breve{n}_B^r, \hat{q}^r) \cup \widetilde{NES}_{i''}^r(\breve{n}_B^r, \hat{q}^r) \tag{3.15}$$

We consider the occurrence possibility of scenario B to be revealed by the simple existence of one candidate $\breve{n}_B^e \in \breve{N}_B^e$ value leading to $\tilde{S}_B^r \neq \{\,\}$. As before, the existence of more than one valid combination leads to a dead-end situation in terms of passive order tracking. The same applies to the cases where despite the existence of a single valid combination, the non-exact liquidity removed solution involves more than one meta-order representing multiple passive orders.

Assuming that all dead-end tracking situations identified through the previous steps have been avoided, zero or one complete solution should have been identified for both scenarios. Therefore, in a first final case where neither scenario A nor scenario B have been validated, we conclude in another dead-end situation that should result from a mismatch between our passive order list and the one prevailing for the actual stock market price level. In the second case where a single scenario is identified as possible, which is the most desirable for order tracking purposes, the complete solution related to the scenario is definitely applied to the price level and the involved orders are saved as executed and/or cancelled. Finally, in a situation where the occurrence of both scenario A and scenario B are considered possible, before concluding in a dead-end situation, we determine if the two complete solutions produce the same results. In such case, the solution is applied to the meta-order list. However, if both complete solutions produce different results, we have to perform our dead-end merging procedure.

As described before, under Rules 2.6 to 2.8 we identify concurrently reported liquidity executed liquidity added events. From an order tracking point of view, we consider Rule 2.6 and Rule 2.8 similar enough to be processed the same way. In these two cases, the price level visible liquidity is completely executed while new liquidity is added. When it comes to Rule 2.8, we ignore the hidden liquidity executed event to focus only on the affected visible liquidity, which should not impact the results. Since we consider all price level passive orders as completely executed before the new liquidity addition, these two cases are handled as a succession of the previous Rule 2.1 and Rule 1.1 processes. First, as seen in Rule 2.1, the meta-orders present in $\tilde{O}_i$ are saved with their execution context information and closed. Then, identical to what was presented regarding Rule 1.1, a new meta-order $\hat{o}^a$ is created to represent the passive orders involved in the added liquidity. It becomes the sole member of the price level meta-order list $\tilde{O}_{i+1} = \{\tilde{o}^a\}$. Since the final price level number of orders and quantity of shares are provided by $\hat{n}^a$ and $\hat{q}^a$, these information are related to the new meta-order properties through (3.3) expressions.

The case reported through Rule 2.7 is more complex since the price level visible liquidity is only partially executed. Back to the events identification procedure, it has been possible to deduce the quantity of shares involved in the liquidity executed and liquidity added events, which are reported through $\hat{q}^e$ and $\hat{q}^a$. However, as claimed before, we initially do not have enough information to make the distinction between the estimated number of totally executed orders $\hat{n}^e$ and the number of orders added on the price level $\hat{n}^a$. From the sole event identification results, we only know these two values to be combined in the price level visible number of orders variation $\Delta n = \hat{n}^a - \hat{n}^e$. Fortunately, under some specific circumstances, it is possible to use our orders tracking concepts to go further and determine these values with our usual level of certainty. In these cases, we are able to identify the passive order(s) affected by the liquidity executed event, as well as characterize the new visible liquidity, which becomes part of the meta-order list.

We begin the process by determining if the characteristics of the executed liquidity may lead to the identification of a liquidity executed exact solution that would also be consistent with the liquidity added event. Similar to function (3.11), the following expression describes this type of potential solution.

(3.16)	$\widetilde{ES}^e = \left\{\tilde{S} \in \mathcal{P}^+\left(\tilde{O}_i\right) \,\middle|\, (3.16.1) \wedge (3.16.2) \wedge (3.16.3)\right\}$

where

(3.16.1)	$\left(\forall \tilde{o} \in \tilde{S}\right)\left(\forall \tilde{o}' \in \tilde{O}_i \backslash \tilde{S}\right)[\tilde{t}_i(\tilde{o}) < \tilde{t}_i(\tilde{o}')]$

(3.16.2)	$\tilde{Q}_i(\tilde{S}) = \hat{q}^e$
(3.16.3)	$\tilde{N}_i(\tilde{S}) + \Delta n > 0.$

It is important to note that no number of orders is involved in the solution $\tilde{S}$ selection which, given proposition (3.16.2), is only exact with regard to the executed quantity of shares. However, this simple match allows us to establish that the number of totally executed orders $\hat{n}^e$ corresponds to the total number of orders of the meta-order(s) included in the solution $\tilde{N}_i(\tilde{S})$. Because of the previously described $\Delta n = \hat{n}^a - \hat{n}^e$ relation, it consequently becomes easy to deduce the number of orders related to the liquidity added event $\hat{n}^a = \tilde{N}_i(\tilde{S}) + \Delta n$, which positivity is ensured by proposition (3.16.3).

Given $\left|\widetilde{ES}^e\right| = 1$, we consider any passive order included solution $\tilde{S} \in \widetilde{ES}^e$ meta-orders as totally executed and we perform the usual actions that consist in closing and saving their contextual information. Denoting the meta-order representing the new liquidity as $\tilde{o}^a$, its post-events properties are provided by the (3.3) expressions. The new composition of the meta-orders set is provided by the following expression.

(3.17)	$\tilde{O}_{i+1} = (\tilde{O}_i \setminus \tilde{S}) \cup \tilde{o}^a$

In the absence of an exact solution, we use the following expression to determine the existence of a non-exact liquidity executed solution allowing for the deduction of $\hat{n}^e$ and $\hat{n}^a$.

(3.18)	$\widetilde{NES}^e = \left\{\tilde{S} \in \mathcal{P}^+(\tilde{O}_i) \,\middle|\, (3.18.1) \wedge \ldots \wedge (3.18.4)\right\}$

where

(3.18.1)	$(\forall \tilde{o} \in \tilde{S})(\forall \tilde{o}' \in \tilde{O}_i \setminus \tilde{S})[\tilde{t}_i(\tilde{o}) < \tilde{t}_i(\tilde{o}')]$
(3.18.2)	$\tilde{Q}_i(\tilde{S}) > \hat{q}^e$
(3.18.3)	$(\forall \tilde{o} \in \tilde{S})[Q_i(\tilde{S} \setminus \tilde{o}) < \hat{q}^e]$

(3.18.4)	$(\exists \tilde{o} \in \tilde{S})\left[\left(\forall \tilde{o}' \in \tilde{S} \setminus \tilde{o}\right)\left[\tilde{t}_i(\tilde{o}) > \tilde{t}_i\left(\tilde{o}'\right)\right]\right.$ $\left.\wedge\left(\begin{pmatrix}\tilde{n}_i(\tilde{o}) = 1 \wedge \\ \Delta n + \tilde{N}_i(\tilde{S} \setminus \tilde{o}) \leq \hat{q}^a\end{pmatrix} \vee \begin{pmatrix}\tilde{n}_i(\tilde{o}) - 1 = -\left(\Delta n + \tilde{N}_i(\tilde{S} \setminus \tilde{o})\right) + 1 \wedge \\ \tilde{n}_i(\tilde{o}) - 1 \leq \hat{q}^e - \tilde{Q}_i(\tilde{S} \setminus \tilde{o})\end{pmatrix}\right)\right]$

As before, a valid non-exact solution $\tilde{S} \in \widetilde{NES}^e$ involves the relation (3.18.4) meta-order $\tilde{o}$ to be considered as partially executed. The remaining meta-orders $\tilde{S} \setminus \tilde{o}$ are considered totally executed. Our goal is to obtain the meta-order $\tilde{o}$ number of totally executed orders $-\Delta\tilde{n}(\tilde{o}) = -(\tilde{n}_{i+1}(\tilde{o}) - \tilde{n}_i(\tilde{o}))$, which then leads to the deduction of the number of incoming orders $\hat{n}^a$. We use the following equation as a stating point :

(3.19)	$\Delta n = \hat{n}^a - \tilde{N}_i(\tilde{S} \setminus \tilde{o}) + \Delta\tilde{n}(\tilde{o})$.

With $\tilde{N}_i(\tilde{S} \setminus \tilde{o})$ corresponding to the solution totally executed meta-orders underlying number of orders, which is known, we rearrange the equation to relate the two unknown values to the known values :

(3.20)	$\Delta n + \tilde{N}_i(\tilde{S} \setminus \tilde{o}) = \hat{n}^a + \Delta\tilde{n}(\tilde{o})$.

Despite their unobserved values, it is possible to establish that in order to be consistent, $\hat{n}^a$ and $\Delta\tilde{n}(\tilde{o})$ must satisfy the following conditions.

(3.21)	$0 \leq -\Delta\tilde{n}(\tilde{o}) \leq \tilde{N}_i(\tilde{o}) - 1$

(3.22)	$\hat{n}^a \geq 1$

Condition (3.21) ensures that at least one passive order, although potentially partially executed in terms of quantity of shares, remains represented by the partially executed meta-order $\tilde{o}$. Since the price level number of orders accounts for a passive order until its last share has been executed, this condition also guarantees that it is possible for $\Delta\tilde{n}(\tilde{o}) =$

0 eventuality. Condition (3.22) simply ensures that it is impossible for a new passive order to enter the price level without a positive effect on the total number of orders.

Based on conditions (3.21) and (3.22), proposition (3.18.4) certifies that a non-exact solution $\tilde{S} \in \widetilde{NES}^e$ belongs in one out of two categories for which a single $\hat{n}^a$ and $\Delta\tilde{n}(\tilde{o})$ combination ensures the consistency of all the involved elements. In the first situation, by combining the fact that $\widetilde{N}_i(\tilde{o}) = 1$ to condition (3.21), we establish that $\Delta\tilde{n}(\tilde{o}) = 0$. Using this information with expression (3.20), it becomes easy to establish that $\hat{n}^a = \Delta n + \widetilde{N}_i(\tilde{S} \setminus \tilde{o})$. In the second case, we use the fact that the $-\Delta\tilde{n}(\tilde{o})$ value cannot exceed $\widetilde{N}_i(\tilde{o}) - 1$ to establish that if the condition $\widetilde{N}_i(\tilde{o}) - 1 = -\left(\Delta n + \widetilde{N}_i(\tilde{S} \setminus \tilde{o})\right) + 1$ is verified, it is possible to claim that $\Delta\tilde{n}(\tilde{o}) = \widetilde{N}_i(\tilde{o}) - 1$ and $\hat{n}^a = 1$. Additionally, in both cases, by combining the already know number of executed shares with the solution totally executed meta-order number of shares, it is possible to establish that the meta-order $\tilde{o}$ number of executed shares $-\Delta\tilde{q}(\tilde{o})$ corresponds to $\hat{q}^e - \widetilde{Q}_i(\tilde{S} \setminus \tilde{o})$.

As before, the totally executed meta-orders information are saved and closed while the partially affected one is updated to reflect the post-events context. We still denote the meta-order representing the new visible liquidity as $\tilde{o}^a$ and use the (3.3) expressions to characterize it with the newly obtained $\hat{n}^a$ and $\hat{q}^a$ values. Then, still assuming that $\tilde{o} \in \tilde{S}$ corresponds to the partially executed meta-order, the following expressions provide its new characteristics as well as the information regarding the price level meta-orders set evolution.

(3.23)	$\tilde{n}_{i+1}(\tilde{o}) = \tilde{n}_i(\tilde{o}) + \Delta\tilde{n}(\tilde{o})$ $\tilde{q}_{i+1}(\tilde{o}) = \tilde{q}_i(\tilde{o}) + \Delta\tilde{q}(\tilde{o})$ $\tilde{O}_{i+1} = (\tilde{O}_i \setminus \tilde{S}) \cup \tilde{o} \cup \tilde{o}^a$

Concluding the complex Rule 2.7 case, it is once again important to note that in the absence of both exact and non-exact solutions, our passive order tracking process reaches a dead end situation and the merging procedure have to be performed to ensure the consistency of future results.

When it comes to Rule 2.9, since the only event affecting the visible liquidity corresponds to the entire price level execution, we directly apply the Rule 2.1 methodology that has been developed for the same purpose.

Finally, as claimed before, events identification Rules 3.1 to 3.4 relate to events sequence that have no visual effect on the price level liquidity. Rule 3.1 have no effect on our passive orders tracking methodology since it only involves the execution of hidden liquidity. On the other hand, in terms of order tracking procedure, Rules 3.2 to 3.4 cases are fortunately identical to their counterparts with visual liquidity effects for which the identified events present the same characteristics. Consequently, no new concepts is required for their treatment. The methodology developed for Rule 2.6 directly apply to Rule 3.2 events, Rule 2.7 to Rule 3.3 and finally, Rule 3.4 is processed using the Rule 2.8 concepts.

3.2 Results

Figure 3.2 presents the general results regarding our passive orders tracking methodology. On a stock basis, this figure reports the rates of visible orders for the four most frequent tracking outcomes. In addition to the proportion of tracked orders ending in identified total execution and cancellation, it presents those related to orders for which, at some point, it has become impossible to keep track and those having moved outside of the 20 price levels window provided by Xetra. Because we lose track of this marginal proportion of orders for mechanical reasons, we do no not explicitly account for them in the tracking success (or failure) rates. In fact, some information regarding these orders remains relevant for future use. Considering the orders belonging in this last category and those whose life-cycle has ended in a cancellation or total execution as successfully tracked, Panels I, II and III show different results regarding the success rate of our methodology according to the stock membership index. Panel I show that for DAX index stocks, which correspond to the 30 largest market capitalizations, the tracking success rates exhibit an important variability. It shows success rates spanning from 16% to 79%, with an average of 41%. Despite these mixed results, it is important to note that regarding these 30 stocks, using our methodology, we have successfully tracked more than 88 millions passive orders over the three month period of interest. A success rate of more than 50% have been achieved for 10 of them.

Panel II shows that when it comes to the MDAX index components, our methodology results become much more interesting. Indeed, for these stocks, the order tracking success rate ranges between 60% and 98%, with an average of 87%. In fact, this rate has been above 90% for half of the 50 represented stocks. Finally, Panel III presents our methodology results regarding the SDAX components, which we qualify as excellent. Ranging between 95.12% and 99.88%, the order tracking success rates present an average of 98.58% for the 50 corresponding stocks.

Since the Figure 3.2 stocks are ordered by number of submitted orders, it is visually obvious that this feature has no real explanatory power regarding the success rate of our tracking methodology. By the nature of this methodology, it became obvious to us that the number of lost passive orders should be related to their cancellation context. Indeed, over the previous section, the ability to keep tracking a price level orders have been intimately related to our ability to retrieve a cancelled order from the list of those elements standing on the price level of interest. It is essentially based on the involved orders quantities of shares. We have also identified order tracking dead-end situations related to the execution of passive orders but, the passive orders execution rate seems too small for these situations to have a considerable impact on the global tracking success rates. Considering the number of orders to choose from in the event of an order cancellation as a possible explanation for a given stock tracking success rate, we have investigated the relation between these two variables. In this context, Figure 3.3 shows a strong log-linear relationship between a stock time-weighted average of the number of orders standing on the five best price levels and our actual tracking success rate. Therefore, to the question of whether it is possible to track orders for a given stock from their submission to total execution or cancellation in a price aggregated limit order book, the short answer is that this capacity is directly related to the usual number of orders present on its book for each price.

3.3 Summary

In this chapter, we have used the Chapter 2 events identification rules to develop an extensive methodology allowing us to track a limit order from its submission to cancellation or total execution.

To the question of whether it is possible to follow the limit orders through a limit order book where the number of orders and quantities are aggregated by price level, we answer that we have been able to perform this task with success rates ranging between 16.4% and 99.9%. For a given stock, we have shown this capacity to be directly related to the average number of standing orders, especially on the first five price levels. We have related the fact that these values are generally increasing for the DAX, MDAX and SDAX components to the 41.5%, 87.4% and 98.6% average tracking success rates obtained using our methodology for these stocks.

Figure 3.1 Liquidity removed passive order tracking

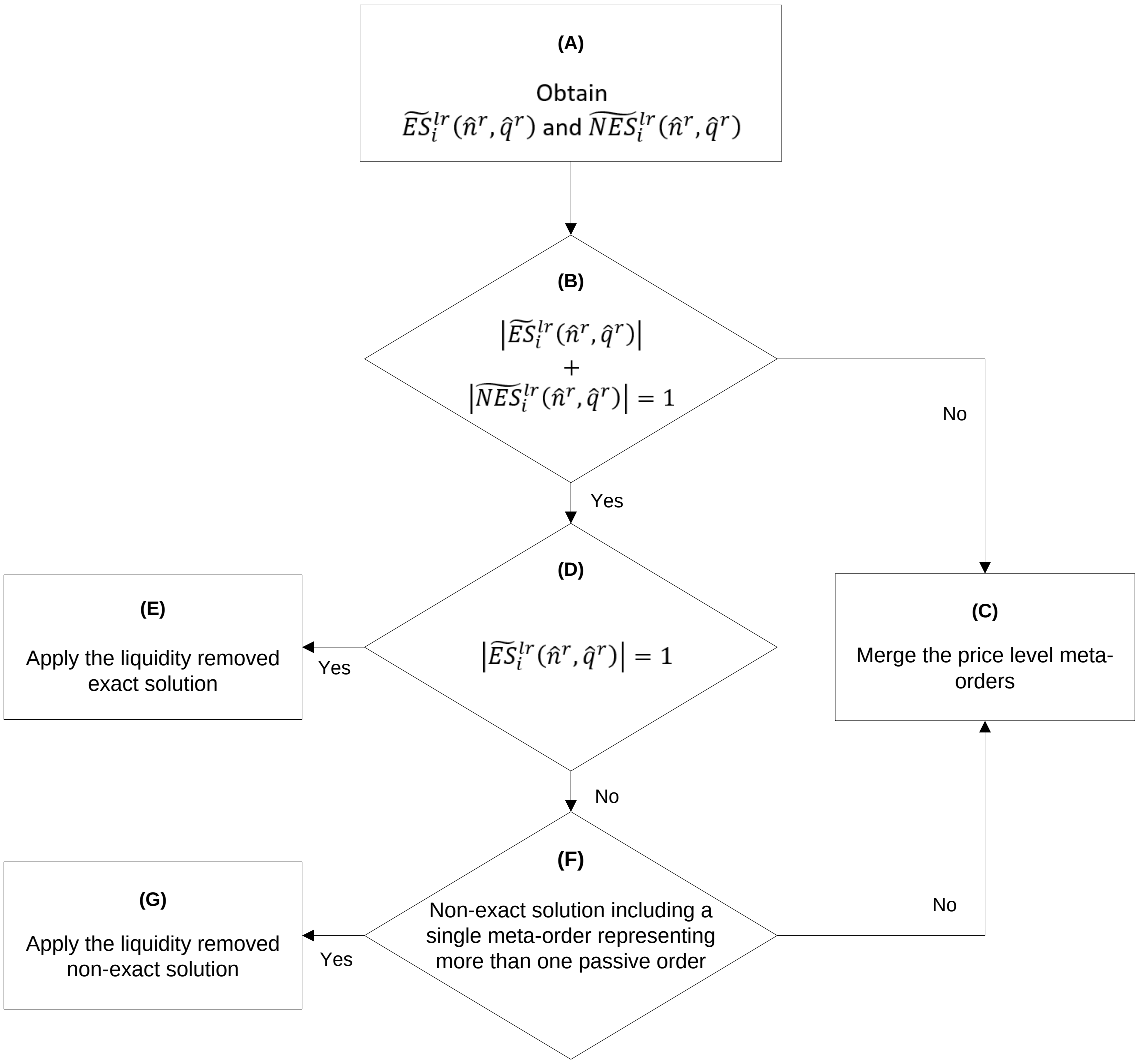
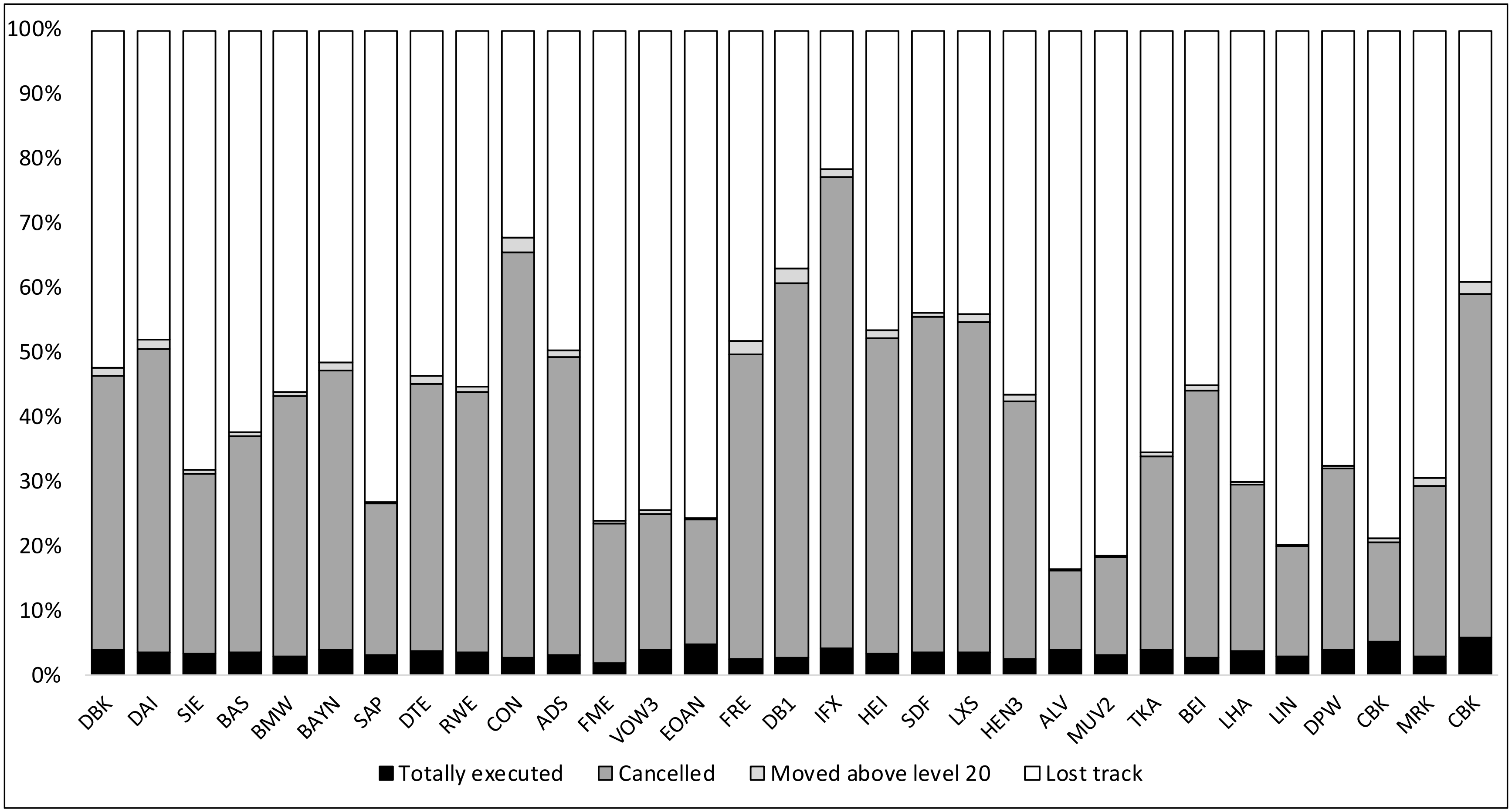
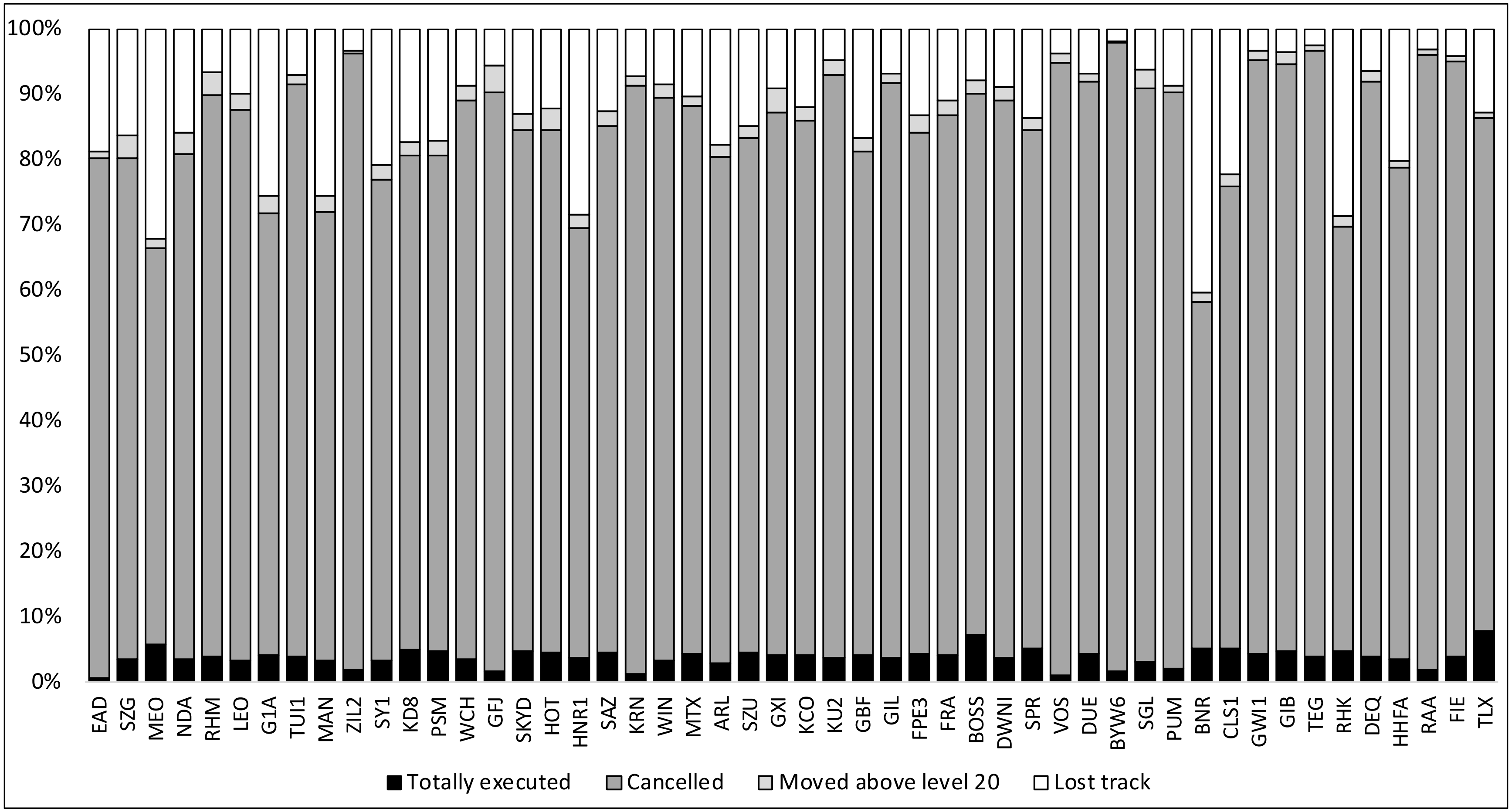


Figure 3.2 Passive orders tracking outcomes

Panel I: DAX index components



Panel II: MDAX index components



Panel III: SDAX index components

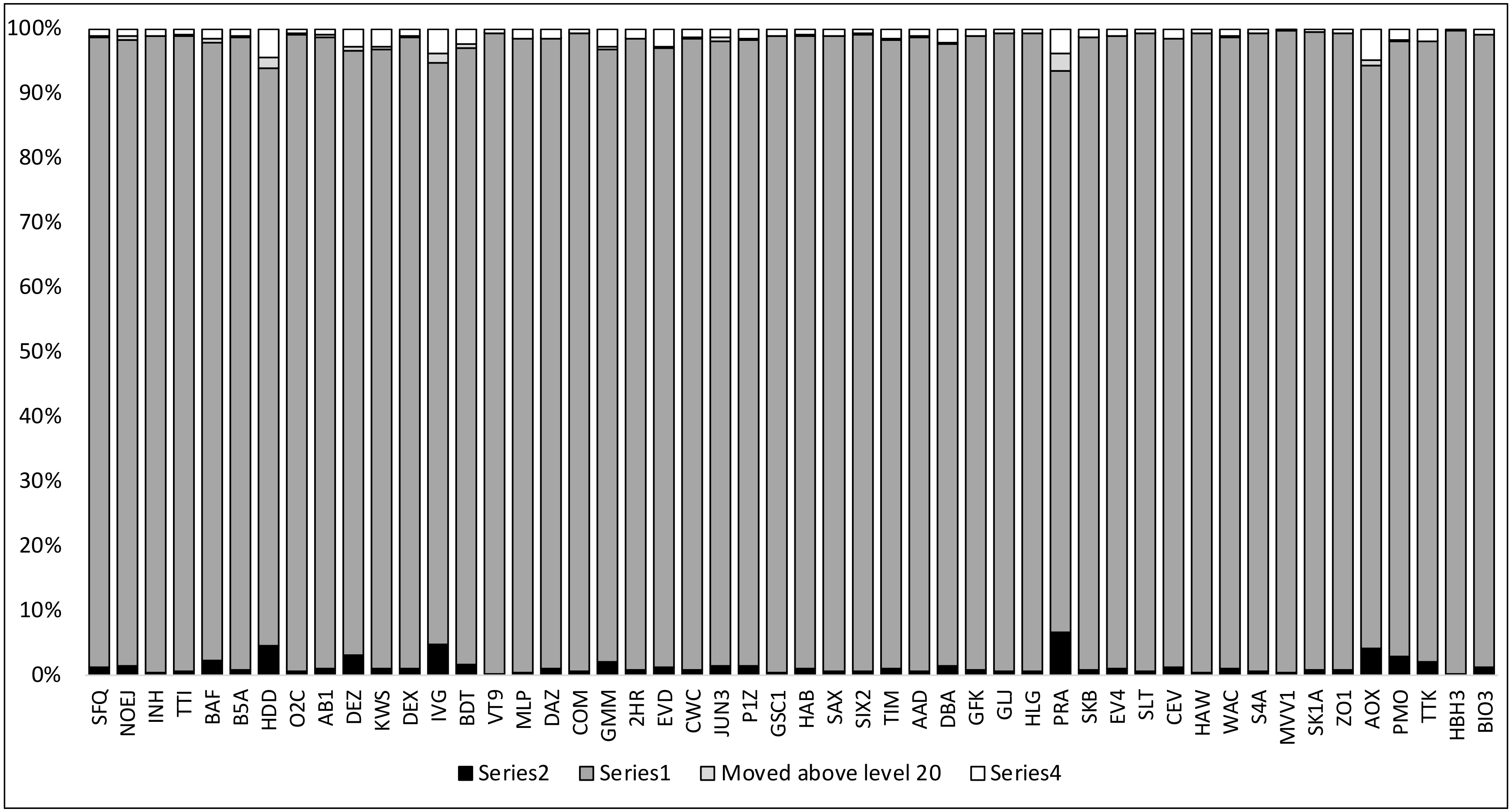
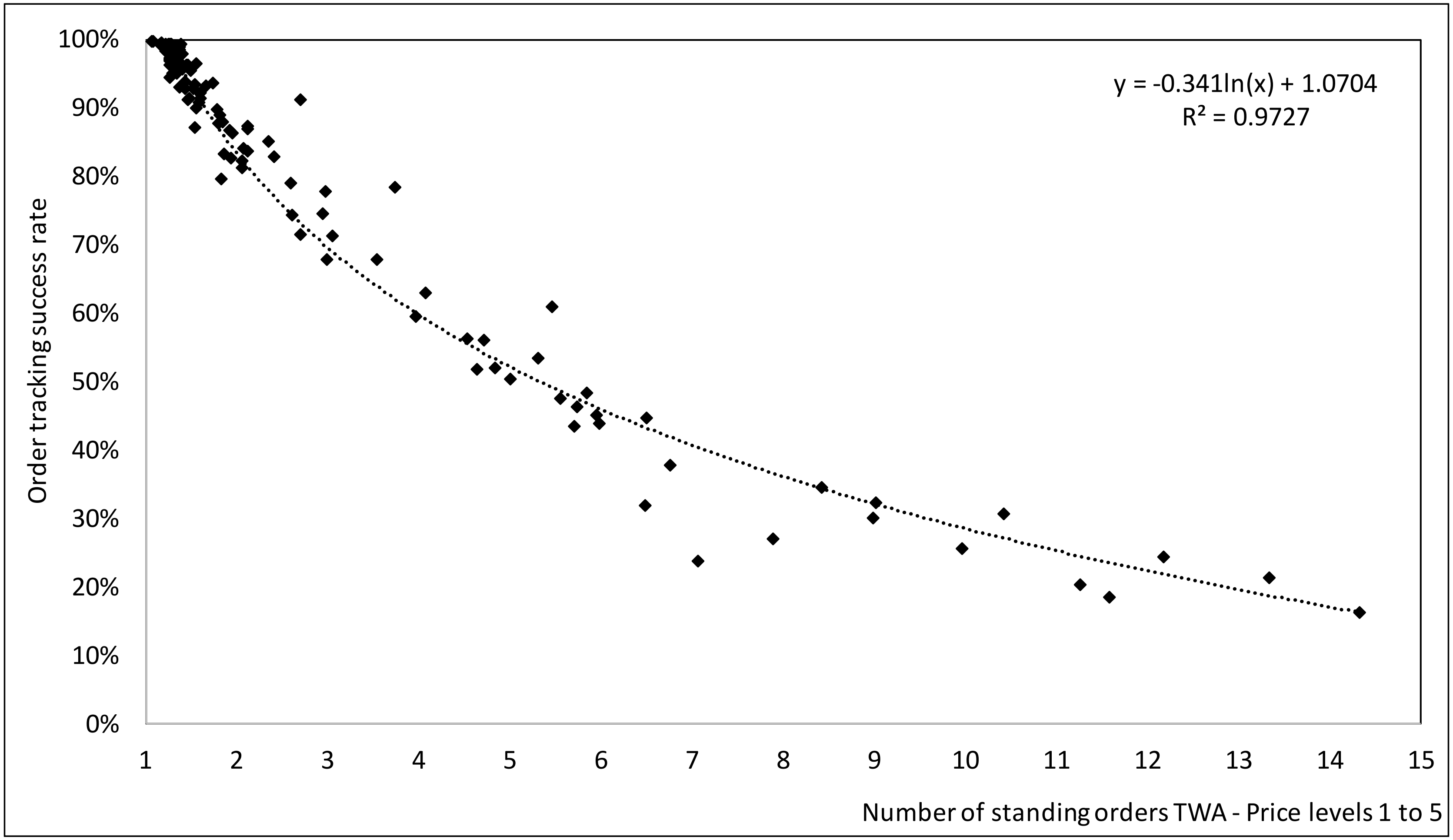


Figure 3.3 Orders tracking success rates



Chapter 4
Deep limit order book events dynamics

In recent years, technology improvements have completely transformed the financial markets landscape. The main part of market activity is now performed through algorithms on electronic trading platforms using real-time open limit order book information. Even more conventional investors such as banks, mutual funds and institutions now outsource their trading tasks to algorithmic traders who split the main orders into multiple child orders distributed over time, trying to achieve the best execution price while hiding their intention to other market participants. In this new financial world, speed and information quality have become very important keys to success. Indeed, in order to remain competitive, algorithms must take investment decisions and send their answer to the market in a few milliseconds time frame and this window size tends to decrease years after years with technological improvements. Although high-frequency and algorithmic trading now represent a market standard around the world, limit order book modeling has not yet received a very important coverage in academic and scientific literature. In this context, this chapter focuses on modeling the behavior of a multilevel limit order book at a microscopic level with algorithmic trading perspectives.

Until recently, regularly spaced data were used for asset returns modeling and forecasting purposes. However, to maximize the advantage of newly available high frequency data, modeling irregularly spaced data is essential since at a microscopic level, important market events such as transaction and limit orders submission are irregularly spaced in time. In this situation, point processes models are natural candidates to describe these irregularities. In a widely cited paper, Cont, Stoikov et al. (2010) use a Poisson process system to represent events arrival in a limit order book and perform various quantities computations, such as the probability of making the spread over a given period. Despite the interesting results produced by their model, the main disadvantage of homogeneous Poisson process remains the assumption of independent exponentially distributed event arrival times. More recently Huang, Lehalle et al. (2015) and Muni Toke and Yoshida (2017) have developed and simulated complete limit order book models based on state-dependant event arrival processes. Although the former introduces a certain level of dependency between processes, none of these papers explicitly consider the arrival of events itself as a major factor driving the arrival of other events. However, it has been established that a clustering phenomenon is present in the arrival sequence of some types of limit order book events. This particularity led to the introduction of self-exciting point processes in finance such as the Hawkes process, that was originally used for earthquakes occurrence modeling and forecasting (Hawkes (1971), Hawkes and Oakes (1974)).

As well described by Shek (2011), in a self-exciting process such as univariate Hawkes model, the current arrival rate for a given type of events is driven by past occurrences of events of the same type. In its multivariate version, Hawkes process is said to be a cross or mutually-exciting process. Within this extended case, in addition to past events of the same type, event arrival intensity is also driven by past events of other types. Using this feature, it becomes possible to account for dependency between different event categories occurring at irregularly spaced times, which has otherwise, not been explicitly covered in the current literature.

Over the last years, Hawkes processes themselves have been used for various purposes in finance. A very general definition of this process and its possible financial applications is provided by Embrechts, Liniger et al. (2011). From another point of view, Bacry, Delattre et al. (2013) use them for price changes modelling purposes. They model the number of ticks up and down for both single and pairs of assets using counting processes. Hawkes processes allow them to take autocorrelation in price movements into account. Fauth (2012) use four bivariate mutually-exciting processes to model bid and ask prices moves up and down. Their processes pairs modeling structure allows them to produce a consistent bid-ask spread. In order to add more information to their model, they also include trades volumes. Large (2007) uses up to ten Hawkes processes to analyse the limit order book resiliency well-known phenomenon.

In a market events perspective closer to the limit order book, Hewlett (2006) uses bivariate Hawkes processes for buy and sell trades occurring on the FX markets modelling. He then applies them for optimal trading strategies purposes as developed by Almgren (2003), Almgren (2000), Bertsimas and Lo (1998), Obizhaeva and Wang (2012) and more recently Cheng, Di Giacinto et al. (2017), assuming the market maker considers buy and sell orders arrival to follow Hawkes processes. Shek (2011) also uses a bivariate mutually-exciting model for buy and sell trades occurrence. As Fauth (2012), he includes trade sizes in order to consider more information in large trades than in small ones. Toke and Pomponio (2012) apply a bivariate process to model the occurrence of trades consuming more than one order book level, also known as trade-through. Finally, Bowsher (2007) uses two pairs of bivariate Hawkes processes to model the joint arrival of trades and mid-quote changes.

In their respective thesis, Vinkovskaya (2014) and Huang (2012) apply multivariate Hawkes processes in order to model the limit order book first level order flows. Vinkovskaya (2014) considers four different processes representing limit order arrivals and market order arrivals for both sides of the book. For estimation procedure simplification, she combines market order and limit order cancellation events. She presents a regime switching extension in order to account for the spread size effect on the order arrival intensities. She performs her parameters estimation using a 2008 subsample of the Trade and Quotes (TAQ) database with a one second time precision. The forecasting power of her model is interesting since it outperforms Poisson model and AR, MA and ARIMA time-series models for out-of-sample predictions. On the other hand, Huang (2012) uses six multivariate Hawkes processes in order to model the arrival of market orders, limit orders and limit order cancellations on both sides of the book. Parameters estimates are obtained from a millisecond precision order book snapshots dataset on a five-day period in 2009 for Vodaphone (VOD.L), a stock traded on the London Stock Exchange. He then computed various probabilities using a Monte Carlo simulation method based on the thinning algorithm introduced by Ogata (1981).

Since only few of the previously cited contributions have analyzed more than one level in the limit order book, we propose to extend the current literature in this way. In this context, we suggest a multivariate Hawkes processes system to model and analyze the behavior of a multilevel limit order book. Our main objective is to determine how various order book event types occurring at different levels affect each other. Let us assume a first level that contain only one limit order and the absence of hidden or iceberg orders. By the nature of the limit order book itself, one can easily see that the cancellation of this order or the submission of a market order consuming it entirely would instantly convert the current level 2 into the new best bid or best ask level. In the same way, level 3 would become level 2 and so on. With this simple example in mind, it is possible to believe that analyzing higher levels event arrival processes could provide interesting information on the future states of best price levels where transactions generally occur. To accomplish this task, we define an extensive set of events occurring on the first twenty levels of both sides of the book and take advantage of the Hawkes dependency structure to establish how they relate to each other.

Our parameters estimation and performance analysis are realized on a Xetra 2013 microsecond (10-6) precision limit order book dataset. It is possible to expect the microsecond precision to provide an informational advantage over what has been done in the past. However, the real Xetra order flows are not directly available and have to be deduced using observable limit order book state changes and executed trades. This procedure may lead to some missing events because of the aggregated nature of the data. On the other hand, our dataset is exactly the one used by algorithmic trading systems that were operating on the Frankfurt Stock Exchange over the February to March 2013 period, suggesting we cannot get closer to real-time market information without having access to the stock exchange internal database. This fact is important since our models are developed in an algorithmic trading perspective. Indeed, we have to keep in mind that it is possible for an algorithmic response to an order book event to be so fast that it could not have been though and launched by a human trader. Fortunately, Hawkes processes have the capacity to capture this type of very short-term phenomena.

The rest of the chapter is organized as follows: In Section 4.1, we present the concept of Hawkes processes, which provides the theoretical foundations for this paper. In Section 4.2, we introduce the dataset. In Section 4.3, we define the set of events considered as potentially interrelated. In Section 4.4, we use Section 4.1 theory to specify our complete model. In Section 4.5, we elaborate the estimation methodology that will lead to one descriptive model for each of our three liquid stocks. In Section 4.6, we analyse the results of the estimation phase in terms of models selection, data fitting and estimated parameters characteristics. In Section 4.7, we use the estimated parameters to describe a global events arrival dynamics that we relate to market participants potential behaviors. Finally, we conclude in Section 4.8.

4.1 Hawkes processes

In this section, we present the theoretical foundations underlying to our models, their estimation methodology, and the analysis performed on the results. Before presenting the fundamentals for univariate and multivariate Hawkes processes models, we introduce the general concepts of point and counting processes.

Carstensen (2010) defines a point process as a *"statistical model used to describe point patterns in a given space"*. In our financial context, since these points generally correspond to events arrival times, this space is one-dimensional and limited to a section of the real line. As reported by Guo and Swishchuk (2020), $(T_1, T_2, T_3, \cdots)$ is called a point process if it is a sequence of non-negative random variables with $P(0 \leq T_1 \leq T_2 \leq T_2 \leq \cdots) = 1$ and if the number of points in a bounded region is almost surely finite. Therefore, the realisation of this point process $(T_1 = t_1, T_2 = t_2, T_3 = t_3, \cdots)$ provides an ordered list $\{t_1, t_2, t_3, \cdots\}$ of time spikes, which corresponds to event occurrence times in this chapter context.

It is convenient to describe a point process through its association with a counting process. Also reported by Guo and Swishchuk (2020), a stochastic process $\{N(t), t \geq 0\}$ is a counting process if it satisfies $N(t) \in \mathbb{N}$, $N(t) \geq 0$, $N(0) = 0$, as well as $\forall t, s \geq 0, N(t+s) \geq N(t)$. It is so called because it actually *counts* the events having occurred at times $\{\{t_i\}, 0 \leq t_i \leq t\}$. Indeed, while $N(t)$ provides the number of events occurring during the time interval $[0, t]$, $N(t+s) - N(t)$ does the same regarding the events taking place on the interval $[t, t+s]$. The trajectories of $N(t)$ are right-continuous and piecewise constant with probability one. Assuming that the times $t \geq 0$ for which the value taken by the counting process $N(t)$ has changed exactly match the events times $\{t_1, t_2, t_3, \cdots\}$ of a point process, the two processes are associated. Therefore, it becomes possible to characterize both processes using the conditional intensity function $\lambda(t)$, which takes the form $\lambda(t) = \lim_{h \to 0} \frac{E\left[N(t+h) - N(t) | \mathcal{F}^N(t)\right]}{h}$, where $\lambda(t) \geq 0$ and $\mathcal{F}^N(t)$ represents the corresponding natural filtration. As an example, the intensity of a pure homogeneous Poisson process, hereafter referred as Poisson process, is provided by $\lambda(t) = \lambda$, a constant parameter corresponding to the average number of event occurrences by unit of time. To avoid confusion, from this point, we will express any point process in terms of its associated counting process.

Having defined the form of the conditional intensity function $\lambda(t)$, it is now possible to define a counting process compensator, which is provided by the following expression :

(4.1)	$$\Lambda(t_0, t_1) = \int_{t_0}^{t_1} \lambda(t)\, \mathrm{dt}.$$

Assuming $t_1 \geq t_0$, $\Lambda(t_0, t_1)$ may be interpreted as the time period t_0 to t_1 expected number of event occurrences.[9] Equation (4.2) provides the general log-likelihood function, which can be maximized to estimate any counting process model parameters.

(4.2)	$$\begin{aligned} \ln\mathcal{L}\left(\{t_i\}_{i=1,\dots,N(T)} | \theta\right) &= \int_0^T \left(1 - \lambda(s|\theta)\right) ds + \int_0^T ln\lambda(s|\theta) dN(S) \\ &= T - \int_0^T \lambda(s|\theta) ds + \sum_{t_i < T} ln\, \lambda(t_i|\theta) \\ &\equiv -\Lambda(0, T|\theta) + \sum_{t_i < T} ln\, \lambda(t_i|\theta) \end{aligned}$$

Univariate Hawkes process

Having defined the general concepts, we now introduce the Hawkes processes models. In its univariate version, such model assumes the occurrence of an event to have an impact on the arrival of events of the same type. Under this paradigm, by its effect on the process intensity function, an event occurrence increases the probability of another occurrence

[9] The Poisson process compensator function is defined as $\Lambda(t_0, t_1) = \int_{t_0}^{t_1} \lambda \mathrm{dt} = \lambda(t_1 - t_0)$

that does the same and so on. Consequently, the process has the potential to become *self-excited*, which makes the description of events clustering phenomena possible.

As defined by Hawkes (1971) and Hawkes and Oakes (1974), in its most general form, the self-exciting Hawkes process is defined as a counting process $\{N(t), t \geq 0,\ N(t) \in \mathbb{N}\}$ characterized by the following continuous intensity function.

(4.3)	$$\lambda(t) = \mu(t) + \int_{s<t} v(t-s)dN(s)$$
where $\lambda(t)$: process intensity $\mu(t) \geq 0$: baseline intensity $v: \mathbb{R} \rightarrow \mathbb{R}^{+}$: excitation kernel function	

The excitation kernel function $v: \mathbb{R} \rightarrow \mathbb{R}^{+}$ represents the residual impact of past events on the current value of the intensity process. To ensure the stability of the self-exciting Hawkes process, this kernel function must meet the $\int_0^\infty v(t)dt < 1$ condition.

Because of discrete nature of $N(t)$, which counts the events taking place at times $\{\{t_i\}, i = 1, \dots, N(t)\}$, it is possible to rewrite (4.3) in the following way :

(4.4)	$$\lambda(t) = \mu(t) + \sum_{t_i<t} v(t-t_i)$$

As it is often the case in financial and economics literature, we use a constant baseline intensity $\mu(t) = \mu$. We also use the simple exponential kernel function, which is characterized by the following expression :

(4.5)	$$v(t) = \alpha e^{-\beta t}$$
where $\alpha > 0$: excitation term $\beta > 0$: exponential decay factor	

Combining (4.5) and (4.4), the univariate exponential Hawkes process intensity function is defined as :

(4.6)	$$\lambda(t) = \mu + \sum_{t_i<t} \alpha e^{-\beta(t-t_i)}$$

In general terms, α represents the immediate impact of an event occurrence on its arrival rate. Each time an event occurs, the process intensity is instantaneously incremented by α. Because of the kernel function, this effect also immediately begins to decline. Considering an event taking place at time t_i, the proportion of α still present in the process intensity at time $t > t_i$ is given by $e^{-\beta(t-t_i)}$. Thus, this quantity depends on both the exponential decay factor β and the time elapsed between t_i and t. It is important to keep in mind that under this exponential framework, an event effect never completely vanishes from the process intensity. However, it remains practically significant over a time span whose length is dictated by β, which brings us to the definition of the half-life period as presented in (4.7).

(4.7)	$$HL = \frac{ln(2)}{\beta}$$

Widely used in biology and physics, this measure is interpreted as the duration over which an effect loses 50% of its strength. It actually loses 99% of its strength after about 6.64 half-lives periods. For computational purposes, we consider it to completely vanish after 20 half-lives periods. Back to (4.6), it is possible to denote that occurrences effects are cumulative. Indeed, $\lambda(t)$ theoretically encompasses the residual effects of all the events having taken place at times $\{t_i\}, i = 1, \dots, N(t)$. The process intensity also include μ, the baseline intensity. This constant represents the event arrival rate in place before the occurrence of the first event. It may also play an important role in the rate prevailing between self-excitation periods when past events effects on $\lambda(t)$ have become marginal.

The following expression defines the branching ratio (BR), which may be interpreted as the average number of child events expected to follow the arrival of a parent event.

(4.8)	$$BR = \int_0^{\infty} \alpha e^{-\beta t} dt = \frac{\alpha}{\beta}$$

This expression leads to condition (4.9), which ensures the stability of process. This condition is intuitive since it is easy to see that a branching ratio equal or larger than one could lead to a process explosion. In such case, it would be expected for each event occurrence to lead to more than one new occurrence.

(4.9)	$$\frac{\alpha}{\beta} < 1$$

The branching ratio is also an important component of the process intensity unconditional expectation, which is provided by the following expression :

(4.10)	$$E[\lambda(t)] = \frac{\mu}{1 - \frac{\alpha}{\beta}}$$

The following expression presents the self-exciting Hawkes process compensator function :

(4.11)	$$\Lambda(s,u) = \int_s^u \mu + \sum_{t_i<u} \alpha e^{-\beta(t-t_i)}\, dt$$

The first term of the integral relates to the baseline intensity while the second cumulates the past events residual effects. Thereafter, given the desirable properties of the exponential Hawkes intensity function, the compensator closed form solution is provided by the following expression :

(4.12)	$$\Lambda(s,u) = \mu(u-s) + \sum_{t_i<s} \frac{\alpha}{\beta}\left[e^{-\beta(s-t_i)} - e^{-\beta(u-t_i)}\right] + \sum_{s \le t_i < u} \frac{\alpha}{\beta}\left[1 - e^{-\beta(u-t_i)}\right]$$

Finally, applying (4.6) and (4.12) to (4.2) leads to the univariate exponential Hawkes process log-likelihood function, which, when computed for the period going from times 0 to T, is provided by the following expression :

(4.13)	$$\begin{aligned}&\ln\mathcal{L}\left(\{t_i\}_{i=1,\dots,N(T)}\right)\\ &= -\mu T - \sum_{i=1}^{N(T)} \frac{\alpha}{\beta}\left[1 - e^{-\beta(T-t_i)}\right]\\ &+ \sum_{i=1}^{N(T)} \ln\left(\mu + \sum_{t_k<t_i} \alpha e^{-\beta(t_i-t_k)}\right)\end{aligned}$$

Since inefficient from a computational point of view, it is transformed into the following expression involving the recursive function $R(i)$, which is appropriate for the purposes of numerical maximization.

(4.14)	$$\ln\mathcal{L}\left(\{t_i\}_{i=1,\dots,N(T)}\right) = -\mu T - \sum_{i=1}^{N(T)} \frac{\alpha}{\beta}\left[1 - e^{-\beta(T-t_i)}\right] + \sum_{i=1}^{N(T)} \ln\left(\mu + \alpha R(i)\right)$$ in which: $$R(i) = e^{-\beta(t_i-t_{i-1})}\left(1 + R(i-1)\right) \forall\, i \geq 2$$ $$R(1) = 0$$

Multivariate Hawkes process

In its multivariate version, the Hawkes process allows to model inter-events arrival dependency. In addition to be potentially affected by the occurrence of events of the same type, arrival of events of a given type can be affected by the occurrence of events of different types. Therefore, in addition to the self-excitation phenomenon described previously, multivariate Hawkes processes allow for mutual-excitation. As described by Embrechts, Liniger et al. (2011), assuming $M \in \mathbb{N}$, a mutually-exciting Hawkes process counting the events taking place at times $\left\{\{t_i^m\}, m = 1, \dots, M, i = 1, \dots, N^m(t)\right\}$ is defined as $\{N^m(t), m = 1, \dots, M, N^m(t) \in \mathbb{N}\}$. As for the univariate case, we focus on the exponential kernel version of the mutually exciting Hawkes process with constant baseline intensity, which is characterized by the following discretized intensity functions.

(4.15)	$$\lambda^m(t) = \mu^m + \sum_{n=1}^{M} \sum_{t_k^n < t} \alpha^{(n,m)} e^{-\beta^{(n,m)}(t-t_k^n)}$$
where $\lambda^m(t)$: component m process intensity $\mu^m \geq 0$: component m intensity baseline t_k^n : time of the k^{th} occurrence of event type n	

In this expression, m corresponds to the type of explained event and $n = 1, \dots, M,$ to those of explanatory events. We refer to n and m as types for the predecessor and successor events. We use the (n, m) notation to designate an events relationship in which event type n acts as the predecessor and event type m as the successor. In such case, we claim event of type n occurrences to have an effect on the occurrence of events of type m. In this context, $\alpha^{(n,m)}$ represents the immediate effect of an event of type n occurrence on type m event intensity. Similarly, $\beta^{(n,m)}$ corresponds to its exponential decay factor. It determines how the impact of past type n events persists in the intensity of type m event over time. It is important to note that while $m \neq n$ represents mutually-exciting relationships, when present, the $m = n$ case relates to a self-exciting relationship. Also, despite the fact that mutual-excitation is supported by the multivariate Hawkes process, dependency structure symmetry is not mandatory. Indeed, assuming $m \neq n$, the presence of a (n, m) relationship does not involve that of an (m, n) relationship.

As useful as in the univariate context, the following expression defines the multivariate Hawkes process compensator :

(4.16)	$\Lambda^m(s,u) = \mu^m(u-s) + \sum_{n=1}^{M} \sum_{t_k^n < s} \frac{\alpha^{(n,m)}}{\beta^{(n,m)}} \left[e^{-\beta^{(n,m)}(s-t_k^n)} - e^{-\beta^{(n,m)}(u-t_k^n)} \right] + \sum_{n=1}^{M} \sum_{s \le t_k^n < u} \frac{\alpha^{(n,m)}}{\beta^{(n,m)}} \left[1 - e^{-\beta^{(n,m)}(u-t_k^n)} \right]$

In the spirit of (4.8), the branching ratio $BR^{(n,m)} = \alpha^{(n,m)}/\beta^{(n,m)}$ is now interpreted as the expected number of type m events related to the occurrence of an event of type n.

As described by Toke (2011), a branching ratios matrix $\mathbf{\Gamma} = \left(\alpha^{(n,m)}/\beta^{(n,m)}\right)_{m,n=1,\ldots,M}$ spectral radius strictly smaller than one represents a sufficient condition for a Hawkes process stability.[10] Afterward, assuming a process stable, the unconditional expectation of its components intensity is given by (4.17), in which $\boldsymbol{\lambda}$ corresponds to the $E[\lambda^m(t)]_{m=1,\ldots,M}$ vector, and $\boldsymbol{\mu}$ to the baseline intensities $(\mu^m)_{m=1,\ldots,M}$ vector.

(4.17)	$\boldsymbol{\lambda} = (\boldsymbol{I} - \mathbf{\Gamma})^{-1}\boldsymbol{\mu}$

The log-likelihood function of a multivariate Hawkes process can be represented as the sum of its components log-likelihood functions. This interesting feature for speed and computational intensity considerations is possible because of the absence of shared parameters between the components. The general process log-likelihood is thus provided by the following expression :

(4.18)	$\ln\mathcal{L}\left(\{t_i^m\}_{m=1,\ldots,M,i=1,\ldots,N^m(T)}\right) = \sum_{m=1}^{M} \ln\mathcal{L}^m\left(\{t_i^m\}_{i=1,\ldots,N^m(T)}\right)$

The log-likelihood function of a single component of the multivariate Hawkes process is provided by the following expression :

[10] The spectral radius of the matrix $\boldsymbol{A}$ is defined as $\rho(\boldsymbol{A}) = \max_{a \in S(A)} |a|$, where $S(\boldsymbol{A})$ corresponds to the set of all eigenvalues of $\boldsymbol{A}$.

(4.19)	$\ln\mathcal{L}^m\left(\{t_i^m\}_{i=1,\ldots,N^m(T)} \middle\| \{t_k^n\}_{n=1,\ldots,M,k=1,\ldots,N^n(T)}\right) = -\mu^m T - \sum_{n=1}^{M}\sum_{k=1}^{N^n(T)} \frac{\alpha^{(n,m)}}{\beta^{(n,m)}}\left[1 - e^{-\beta^{(n,m)}(T-t_k^n)}\right] + \sum_{i=1}^{N^m(T)} ln\left(\mu^m + \sum_{n=1}^{M}\sum_{t_k^n < t_i^m} \alpha^{(n,m)} e^{-\beta^{(n,m)}(t_i^m - t_k^n)}\right)$

Finally, in the way (4.14) did before, (4.20) provides a more computationally efficient version of this function.

(4.20)	$\ln\mathcal{L}^m\left(\{t_i^m\}_{i=1,\ldots,N^m(T)} \middle\| \{t_k^n\}_{n=1,\ldots,M,k=1,\ldots,N^n(T)}\right) = -\mu^m T - \sum_{n=1}^{M}\sum_{k=1}^{N^n(T)} \frac{\alpha^{(n,m)}}{\beta^{(n,m)}}\left[1 - e^{-\beta^{(n,m)}(T-t_k^n)}\right] + \sum_{i=1}^{N^m(T)} ln\left(\mu^m + \sum_{n=1}^{M}\sum_{t_k^n < t_i^m} \alpha^{(n,m)} R^{(n,m)}(i)\right)$ in which, $\forall\, i \geq 2$ $R^{(n,m)}(i) = e^{-\beta^{(n,m)}(t_i^m - t_{i-1}^m)} R^{(n,m)}(i-1) + \sum_{t_{i-1}^m \leq t_k^n < t_i^m} e^{-\beta^{(n,m)}(t_i^m - t_k^n)}$ and $R^{(n,m)}(1) = \begin{cases} \sum_{t_k^n < t_i^m} e^{-\beta^{(n,m)}(t_i^m - t_k^n)}, & N^n(t_i^{m-}) > 0 \\ 0, & N^n(t_i^m) = 0 \end{cases}$

4.2 The data

As introduced before, this chapter empirical analysis is performed using our extensive three months Xetra dataset. Given the important number of parameters that may result from our models selection and estimation methodology, we restrict our analysis to the stocks of three individual companies : BMW, SAP, and Adidas. This restriction increases our high-dimensional result sets intelligibility. BMW is an automobiles and motorcycles manufacturer, SAP an enterprise software corporation and Adidas (ADS), an important player of the shoes, clothing, and accessories industry. These companies stocks are DAX index components, which encompasses the 30 main German blue chips trading on the Frankfurt Stock Exchange.

As before, the time period covered in this chapter spans from February 1st to March 31st, 2013. It encompasses 61 trading days over which BMW, SAP and ADS stand in the DAX Index second tier in terms of the total traded volume. These *liquid stocks* respectively occupy 16^{th,} 12th and 19th ranks with approximately 115, 187 and 58 million of traded shares. These ranks slightly improve when it comes to the total traded amount. Regarding this second metric, our stocks occupy the 11th, 7^{th,} and 14th positions with rough totals of 8, 11.4 and 4.4 billions EUR. Similar to traded shares, although close in ranking, SAP exchanged amount is 42.5% larger than BMW that already represents an 81.8% improvement over ADS. This documents the fact that when it comes to traded volumes, the DAX index components show important discrepancies. As an example, during our reference period, the average exchanged amount of stocks constituting the most traded half of the DAX index in these terms appears 3.3 times larger than for those included in the other half. Therefore, we claim that selecting our three stocks inside the second tier of the DAX index in terms of traded volumes ensures diversity even if they seem closely related at some levels. From this point, to simplify the reading of information related to our three liquid stocks, we use a vector inspired notation in which [BMW; SAP; ADS] reports the values specific to BMW, SAP, and ADS.

A DAX stock normal trading day essentially consists in five steps. The opening auction takes place between 8:50 and 9:00.[11] It is followed by a four-hour continuous trading session that is suspended at 13:00, leaving the room for the intraday auction that last for a minimum of 2 minutes. Afterward, the market resumes for a four and a half hour continuous trading session ending at 17:30, which corresponds to the closing auction start time. This last auction period ends between 17:35:00 and 17:35:30. In order to focus on the most typical intraday market conditions, we discard the post opening auction as long as the pre and post intraday auction 30 minutes continuous trading periods. As a result of these exclusions, our trading day consists in both a morning and an afternoon period going from 9:30 to 12:30 and 13:30 to 17:30, which sums to 7 trading hours.

Although transformed, our dataset originates from Xetra Enhanced Broadcast Solution 13.0, a piece of software responsible for transmitting real-time market information to Frankfurt Stock Exchange participants. In addition to executed transactions information, it includes microsecond timestamped update information for 20 price levels on both sides of the limit order book (LOB). Regarding our 61 seven-hour trading periods, BMW, SAP and ADS approximately count 114, 106 and 72 million of these LOB updates. It is interesting to notice that although less traded in terms of shares and euros volumes, BMW is slightly more active than SAP from a LOB perspective. An update is reported for any change occurring on the first 20 price levels of both sides of the Limit Order Book (LOB). They are generally related to trade executions, limit order submissions, and limit order cancellations. It is possible for multiple price levels variations to be reported through the same update. Figure 4.1 presents the empirical cumulative distribution functions of the durations between reported LOB updates. With the exceptions of the proportion of changes being reported in less than one microsecond that appears to be more important for BMW and those with an interval between 10 and 160ms that seems slightly higher for SAP, this figure shows that our stocks present similar characteristics in terms of LOB update durations. More than 75% of these variations are separated by less than 100 ms

[11] Opening, intraday and closing auction sessions have a random ending time occurring during the 30 second period following 9:00:00, 13:02:00 and 17:35:00.

and about 95% of them by less than one second. These large proportions of close LOB changes are consistent with the presence of high frequency trading activities.

4.3 The events

Our first step in describing a multi-level order book events arrival dynamics consists in identifying the events of interest. At the highest level, market participants interact through orders submission and cancellation. We consider each of these individual actions as an event occurrence that must be classified based on its effect on the LOB. As long as their impact on available liquidity, best available prices and bid-ask spread, we use the affected LOB side and depth level as events classification drivers. We use the events identified using the Chapter 2 methodology as our main data source. From our financial point of view, we also take the investor intention behind each event type into consideration. As an example, assuming they involve the same number of shares, a trade execution and a limit order cancellation may have the same effect on the best available price and available liquidity. However, the intentions behind the decisions leading to the new order book state being clearly different, we consider important to make a distinction between these two types of events. Widely inspired by Large (2007), we divide the events into four global categories : Trades with best price impact (Trades w/ BPI), Trades without best price impact (Trades w/o BPI), Limit order submissions with best price impact (LOS w/ BPI) and LOB events. We divide this last category into limit order submissions taking place on the current best price depth level and beyond (LOS), and Limit order cancellations (LOC).

An actual share exchange taking place on an order-driven stock market generally results from the submission of a market order or an aggressive limit order. We define this second element as a bid (ask) limit order for which the price is equal or higher (lower) than the best available ask (bid) price, which lead to its immediate partial or total execution. Our dataset providing no explicit information on these transactions starting points, depending on the affected book side, we simply refer to these events as buy and sell trades. However, in our attempt to capture the largest possible number of dependency effects, we follow Large (2007) and make the distinction between trades affecting the best available price and the bid-ask spread from those who do not. In our Xetra context, we relate these categories to transactions entirely consuming the first depth level and beyond, or trades-through (see Toke and Pomponio (2012)), and those partially affecting this best price level. We denote Trades w/ BPI as Buy* (Sell*) events and Trades w/o BPI as Buy (Sell). Panel I of Figure 4.2 visually presents a Buy* trade event example. The best ask price level being entirely consumed, while the bid-ask spread increases from 0.02 to 0.03, the pre-trade second best ask price becomes the post-trade best ask price. Similarly, Panel II shows the effects of a Buy event. In this case, the best ask level is partially consumed with no effect on the best ask price and the bid-ask spread.

We define a distinct event categories for LOS and LOC events taking place on the bid and ask sides of the book on a depth level basis. As with trade events, we make the distinction between LOS with and without best price impact. While referring to bid and ask LOS w/o BPI events as BA1 and AA1, we identify bid and ask LOS w/ BPI as BA1* and AA1*. Representing the arrival of a new limit order inside the bid-ask spread, these last events present some particularities. First, for these events to be feasible, the spread must be wide enough to allow the creation of a new price level. Consequently, in situations where the bid-ask spread is only one tick wide, it is impossible for BA1* and AA1* events to be observed. Because of this specificity, Zheng et al. (2014) have modeled these events arrival using a constrained process. However, our BMW, SAP and ADS stocks presenting a bid-ask spread wider than one tick for 92, 70 and 88 percent of the complete 61 trading days dataset, we keep relying on the traditional multivariate Hawkes processes for the representation of each of our events arrival sequences. Second, assuming the creation of a price level inside a favorable spread, the only event occurrence identified as BA1* or AA1* is the one related to the submission of this new level first order. As long as its price remains the best available, depending on its book side, any LOS taking place on this depth level will be identified as a BA1 or AA1 event. Panel III of Figure 4.2 provides a simple ask LOS w/ BPI event example. As for the previous examples, before the event occurrence, the best available ask price is 1.06 with a 0.02 bid-ask spread. Then, an ask 1.05 limit order is submitted, leading to the creation of a new best ask price level and a decreased bid-ask spread that goes from 0.02 to 0.01, which situation is identified as the occurrence of an AA1* event.

Our extension of Large (2007) events classification mostly concerns limit order submissions without best price effect and limit order cancellations. In this paper, the author defines four event types, one for each combination of book side (bid or ask), and LO action (submissions or cancellations). In order to include a depth dimension into our LOB events arrival analysis, in addition to the book side, we group LOS and LOC events on the basis of the depth level number, which is relative to the best bid or ask price on which they take place. Disregarding the involved prices, we refer to LOS w/o BPI events taking place on the twenty referenced depth levels as BA1 to BA20 for the bid book side, and AA1 to AA20 for the ask side. Similarly, we denote bid and ask LOC events as BC1 to BC20 and AC1 to AC20. This is to reduce the number of categories required to cover the same price range that we choose this best prices relative approach rather than that of Cont et al. (2010) who uses prices ticks grids in order to define their events arrival processes. Because of this choice, we identify three possible scenarios leading to the identification of the same LOS event type. Figure 4.2 provides an example for each of them where the type of the identified event is AA2. First, Panel IV presents the trivial case where a limit order is submitted on the existing second best ask price level, increasing its available liquidity. Second, in the Panel VI example, no liquidity with a 1.07 price is available in the pre-event LOB. Therefore, resulting from a limit order submission leading to the creation of a new depth level on this available tick space, the event is identified as taking place on the second best price level, despite the pre-event prevalence of a different depth level 2. With the same idea, Panel VI presents a situation where a submitted limit order also leads to the identification of a depth level 2 LOS event despite the fact that room is available for a 1.07 new second best price level. In this case, the post-event 1.08 involved price level still corresponds to the second best price. Closing our events identification examples, Panel VII finally presents a BC1 event case in which the liquidity available on the bid best price level is decreased by the size of the cancelled order.

Having formally identified the events of interest, Figure 4.3 presents daily average number of the events studied in this chapter for our three liquid stocks. Globally, Panels (a) to (c) suggest limit order submissions to be more frequent than transactions. Disregarding their detailed classification, for each trade event, our liquid stocks are subject to an averages 30.5, 26.9 and 28.2 limit order submissions. When it comes to the most aggressive events, Panel (a) shows Trades w/ BPI to be slightly more frequent than Trades w/o BPI. The actual number of shares consumed in these trades being not considered in our classification, it is not surprising to observe an almost equal numbers of Trades w/ BPI and Trades w/o BPI events. Although likely for very large trades to consume more than one price level, these events classification remains dependent of the LOB shape. For the same number of consumed shares, it is possible for a trade to be classified as either Buy* (Sell*) or Buy (Sell), depending on the pre-trade LOB. As we decrease on the events underlying action aggressivity scale, Panels (b) and (c) suggest that while more common than Trades, BA1* and AA1* appear less frequent than BA1 and AA1. We relate this observation to the fact that LOS w/o BPI events are not restricted by the previously described bid-ask spread conditions and involve less commitment than LOS w/ BPI underlying actions. Finally, regarding deeper LOB events, Panels (b) to (d) show that for LOS as well as LOC, an important part of the action actually take place on depth levels 1 to 5. Beyond this point, we denote a decrease in events arrival up to depth level 10 for LOS and level 11 for LOC where some local peaks are observed. Afterward, the number of occurrences appears generally constant up to level 17. Finishing with depth levels 18 to 20, although LOC events appear slightly more common than LOC, a very small number of both event types appear to take place this far from the best bid and ask.

4.4 The model

In this section, we relate Hawkes processes theory to our events definition through the definition of our potential arrival processes. We begin by defining the set of all events using the following expression :

(4.21)	$S = \left\{ \begin{matrix} Buy^*, Sell^*, Buy, Sell, BA1^*, AA1^*, \\ BA1, \dots, BA20, AA1, \dots, AA20, BC1, \dots, BC20, AC1, \dots, AC20 \end{matrix} \right\}.$

In addition to Trades w/ BPI, Trades w/o BPI and LOS w/ BPI, this set includes LOS w/o BPI and LOC events taking place on depth levels 1 to 20 for both sides of the book, which represents a total of 86 event types. Obtained from equation (4.15), the next expression represents the intensity function of any event $m \in S$ process on a given trading day $d = 1, \dots, 61$.

(4.22)	$$\lambda_d^m(t) = \mu_d^m + \sum_{n \in S^m} \sum_{t_k^n < t} \alpha_d^{(n,m)} e^{-\beta_d^{(n,m)}(t - t_k^n)}$$

In this definition, S^m corresponds to the set of predecessor events used to explain the arrival of the successor event m. In a complete model context, $\forall a \in S, S^a = S$, which is related to the previous section notation by $|S| = M$. However, S representing 86 event types, a complete model would involve 7 396 events relationships, which would lead to the estimation of 14 878 parameters on a daily basis. We establish rules to limit the number of explanatory events involved in some of the intensity processes. Based on events categories and affected depth levels, we define four sets of explanatory events acting as starting points in the estimation methodology presented in the next section. These definitions are presented in Table 4.1. Given a successor event $m \in S$ (column), S^m includes the checked explanatory events groups (rows). First, Trades and LOS w/ BPI arrival intensity processes are initially unrestricted. Indeed, taking advantage of the data availability, we test these successor events arrival for all potential dependency relationships. Second, Trades and LOS w/ BPI are part of the potential explanatory events set of all LOB events. Then, we define the exact set of each of them on the basis of its occurrence depth level. Following an incremental pattern, the potential explanatory events sets for LOS and LOC events taking place on depth levels 1 to 3, 4 to 6 and 7 to 10 respectively include levels 1 to 3, 1 to 6 and 1 to 10 LOB events. In a last increment, no restriction are imposed on depth levels 11 to 20 LOB events arrival intensity processes. By imposing these restrictions, the number of relationships potentially included in our daily models is reduced to 5268, which leads to 10622 estimated parameters. We consider this 29% reduction as an interesting trade-off between model exhaustivity and estimation efficiency.

4.5 Estimation methodology

Having defined an extensive set of potentially related events and imposed initial restrictions to our multivariate Hawkes process, we intent to identify the most recurrent dependency relationships observable for our three stocks over our data sample period. Each trading session being unique in many ways, we do not attempt to analyze each daily idiosyncrasy present in these stocks events dependencies structures. However, by focussing on the significant and persistent relationships, we expect to identify and depict some of their more general characteristics. In order to identify these interrelations, we develop a two part estimation methodology that we apply on each of our three liquid stocks. First, on a trading day basis, we identify the members of the previously defined sets of potential explanatory events that have a significant effect on the arrival of their respective successor event. We refer to the daily Hawkes processes made of these relationships as *complete models*. Then, selecting the most recurrent relationships from these daily models, we define a final Hawkes process that we identify as a *descriptive model*. Afterward, it becomes possible to perform our analysis using the daily parameters estimated value of the three descriptive models. It is important to note that despite the fact that the initial sets of potential explanatory events are the same for each stock, since they do not present the same recurrent events relationships, we observe discrepancies across the components of the three descriptive models.

Estimating high-dimensional Hawkes processes parameters presents some challenges. The log-likelihood function being not strictly concave, it is possible for the numerical maximization procedure to find a local maximum instead of a global one.[12] Our experiments suggest this method to offer a better performance in situations where the underlying process is not crowed by multiple non-significant relationship parameters to estimate. Parameters initial value selection also appears to represent a key factor to convergence. Having these issues in mind, we establish that in our high-dimensional context, for a given successor event, estimating the parameters of a Hawkes process simultaneously by initially including all the concerned predecessor events may not represents the best avenue. Indeed, our main objective being to identify relevant (predecessor, successor) events relations without prior assumption, testing all of them in a single pass may represent an attempt to estimate multiple non-significant parameters. In the same spirit, providing appropriate initial values for all the estimated parameters may

[12] We perform parameters estimation using Matlab interior-point algorithm. In situations where this algorithm does not converge correctly, the BOBYQA algorithm implemented in the free NLOpt package is used for a second estimation pass. In the rare cases for which this second algorithm also fails to converge, the tested dependency relationship is discarded.

become a challenging task. In this context, we use an iterative methodology allowing to divide the whole parameters estimation task into smaller sub-tasks. By successively introducing potential relationships into the model, it becomes possible to discard the less relevant ones early in the process.

Working on a stock, trading day and dependant event basis, our methodology first step consists in estimating the parameters of a simple univariate Hawkes model for each potential explanatory event. Each of these models assumes the included predecessor event type as the only one involved in the successor event intensity process. The main objective of this first step is to obtain parameters initial values to be used in the next estimation round in which candidate explanatory events are successively introduced into the model. It also allows us to go through first relationships discarding round. Explanatory events for which the estimated α parameter p-value is larger than 0.01 are immediately excluded. These values are obtained by inverting the Hessian matrix at the log-likelihood maximum point identified by the non-linear optimization algorithm. We also exclude events relationships for which the α parameter value, although appearing significant, is smaller than 0.1. Finally, we discard any relationship presenting a branching ratio (BR) smaller than 0.02. In this case, we consider the predecessor event to have a neglectable effect on the successor event arrival. By performing this first discarding task early, we attempt to ease the next estimation rounds that may benefit from reduced sets of potential explanatory events.

The second step represents the core of our estimation methodology. It generally consists in multiple model estimations for each explained event type. For each of them, we begin with the set of predecessor events that have not been discarded from the previous step. We build the foundations of each daily complete model iteratively. On each iteration, a new explanatory event candidate is introduced into the evolving model. We use parameters value obtained in the first step as optimization algorithm initial values for the newly introduced predecessor event. Once the model parameters are estimated, events relationships appearing non-significant or weak are discarded. We keep using the criteria described in the first step to perform these exclusions. The non-discarded explanatory events remain in the model for the next iteration. In these cases, rather than using step one coefficients as initial values for the optimization algorithm, we use their last estimated values. This iterative procedure goes on until all step one predecessor events candidates have been tested in the model. At this stage, we identify the Hawkes process model containing the non-discarded events as the *complete model* for the trading day.

For each stock, the last step of our methodology consists in building a *descriptive model* and estimating its parameters value for each trading day. We simply use the number of days for which event relationships have been confirmed as a part of a daily complete model to select the components of these final models. We choose each predecessor event considered as having a significant effect on the successor event on at least 31 days. This threshold ensures the relationships included in our final models to have been part of more than 50 percent of the daily complete models. Once this task is completed, each model parameters are estimated on a daily basis. These estimated values that represent a source of events dependency information are used to perform our next sections analysis.

4.6 The results

4.6.1 Descriptive models selection

Reporting results from the first step of our estimation methodology, Table 4.2 presents the number of trading days for which events relationships meet our selection criteria. Despite relying on preliminary observations, we consider this table as a general events arrival *dependency map*. The different panels split explained event types by categories and depth levels of occurrence. Each cell represents a potential relationship for which the row and the column link a predecessor event to a successor event. We use the number of trading days over which the selection criteria are met to classify these relationships. First, we qualify events effects encountering the selection conditions on 0 to 6 trading days (out of 61) as *absent*. Conversely, we define those for which the numbers of days lies between 55 and 61 as *persistent*. While fairly straightforward to categorize these first relationships groups, the task increases in ambiguity when it comes to those presenting a number of effective trading days lying between these two opposite ranges bounds. In this context, we define events effects meeting our significance criteria on 7 to 30 trading days as *sparse*. We relate these effects to algorithms and trading mechanisms that have a significant impact on a limited number of trading days. On the other hand, because of their impact on the main part of our analysed period, we classify effects observed on 31 to 54 days as *frequent*. Therefore, having established our selection threshold to 50% of the trading days, these last effects combine to those identified as *persistent* to form our descriptive models dependency components.

Table 4.2 Panel I first shows that for our liquid stocks, Trades w/ BPI and Trades w/o BPI events appear only affected by Trades w/o BPI and LOS w/ BPI on a recurrent enough basis to be included in our descriptive models. As mentioned before, we have defined our initial candidate models in order to be able to take advantage of the data availability and analyse the effects of LOB events on the arrival of Trades w/ BPI and Trades w/o BPI events. However, having performed an extensive search for the presence of BA1 to BA20, AA1 to AA20, BC1 to BC20 and AC1 to AC20 effects on the arrival of Buy*, Sell*, Buy and Sell events, we conclude in the general absence of such relationships. As presented in Panel I, despite some sparse exceptions, most of these relationships have met our selection criteria on less than 7 trading days, which are qualified as absent based on our previously defined scale. We relate this lack of recurrent effects to the proportion of LOB events relative to Trades. In line with the previously presented numbers, regardless of the depth level on which they take place, disregarding LOS w/ BPI, the average number of LOB events appears 50.8 to 57.3 times higher than the number of Trade events. This information suggests the LOB to be the theatre of multiple LOS and LOC strategies that, most of the time, do not lead to actual shares exchange. Therefore, we consider consistent that the Trades events arrival do not appear systematically affected by single events arising from these extensive games. On the other hand, the presence or absence of Trades w/o BPI and LOS w/o BPI effects on Trades w/ BPI and Trades w/o BPI events seems related to the LOB side affected by both predecessor and successor events. Indeed, Panel I shows that while (Buy, Buy*), (Sell, Sell*), (Buy, Buy) and (Sell, Sell) appear frequent to persistent, it is possible to qualify (Buy, Sell*), (Sell, Buy*), (Buy, Sell) and (Sell, Buy) as absent. Then, presenting lower contrast levels, (AA1*, Buy*) and (BA1*, Sell*) appear to be highly persistent for our three stocks while (BA1*, Buy*) and (AA1*, Sell*) span from sparse for SAP, to frequent in the cases of BMW and ADS. Consequently, the SAP descriptive model imposes restrictions on these two last relationships, which is not the case for BMW and ADS. We use the same guidelines for (AA1*, Buy) and (BA1*, Sell) that appear frequent in the BMW and SAP cases but meet our selection criteria on less than half the trading days for ADS.

Panel I of Table 4.2 also indicates that LOS w/ BPI (BA1* and AA1*) arrival seems mainly affected by Trades and other BA1* and AA1* events. With some rare sparse exceptions, these events appear generally unaffected by LOS and LOC occurring on depth levels 2 to 20. The situation differs when it comes to LOS and LOC events taking place on the prevalent best price levels. We observe some BA1, AA1, BC1 and AC1 events effects on the arrival processes of BA1* and AA1*. The actual presence or absence of such relationships appear directly related to the involved events book side of occurrence. Indeed, in a radical contrast with the high (BA1, BA1*) and (AA1, AA1*) persistence levels, we note a complete absence of (BA1, AA1*) and (AA1, BA1*) relationships. Similarly, while (BC1, AA1*) and (AC1, BA1*) appear highly persistent, we observe various levels of recurrence for (BC1, BA1*) and (AC1, AA1*) that go from sparse to persistent. In this context, since it meets our criteria on an insufficient number of trading days, (AC1, AA1*) is excluded from our BMW descriptive model. Trades w/o BPI effects on LOS w/ BPI also appear related to each event book side. Indeed, while (Buy, BA1*) and (Sell, AA1*) show an important level of persistence, (Sell, BA1*) and (Buy, AA1*) appear almost inexistent. Finally, the persistent (Buy*, BA1*) and (Sell*, AA1*) and frequent to persistent (Sell*, BA1*) and (Buy*, AA1*) relationships preliminarily suggest LOS w/ BPI events occurrence probability to be increased by the two types of Trades w/ BPI events.

Representing the last elements of this preliminary results overview, Panels II to IX of Table 4.2 present relationships involving LOS w/o BPI and LOC events as the successor counterpart. Since most of these results appear symmetrical for both sides of the book, we focus on dependent events taking place on the bid side of the book, which correspond to Panels II to V. These panels show that the recurrence level of some relationships sharing predecessor and successor events category may appear related to the depth levels over which these involved events take place. This translates into *dependency zones* delimited by predecessor and successor events depth level of occurrence over which, relationships may present increasing, decreasing, or similar levels of recurrence.

As shown in Panels II to V, the number of event types affecting LOS w/o BPI and LOC arrival recurrently enough to be included in our descriptive models appears non monotonically decreasing as the depth level on which these successor events take place increases. For our liquid stocks, while presenting a peak regarding BA1 to BA4 and BC1 to BC4, the trend becomes generally decreasing when it comes to LOB events occurring on depth levels 5 to 12. Beyond this point, when greater than zero, the number of event types having a recurrent effect on the arrival of LOS w/o BPI and LOC is limited to one or two. Panels III and V show that depending on their category, successor events appear recurrently affected by the arrival of at least one predecessor event type up to a depth level rank lying between 16 and 19. BA11 to BA20 and BC11 to BC20 events arrival processes appear generally unaffected by other LOS w/o BPI and LOC occurrence, no matter the depth level on which these predecessor events take place. Still shown in Panels III and V, with three frequent and one persistent exceptions regarding LOS w/o BPI, the 1600 relationships of these natures that have been processed through our estimation methodology appear at most sparse. We observe important dependency zones over which relationships have not met our selection criteria on a single trading day.

Panels II to V also suggest LOS w/o BPI and LOS arrival to be generally affected by Trades and LOS w/ BPI events. We observe some of these events to have a recurrent effect on LOS w/o BPI arrival up to level 16 and LOC up to level 19. With rare exceptions, they remain the only events identified by our methodology as having a recurrent effect on LOB events taking place above level 11. Leaving the detailed dynamics description to a subsequent section, we observe the absence, presence, and recurrence of these relationships to appear highly related to the LOB side affected by the involved events. Moreover, although presenting different patterns, the number of trading days over which relationships meet our selection criteria appear decreasing as we move higher in the depth levels. Taking the example of (Trades w/ BIP, LOS w/o BPI), despite some minor discrepancies among our liquid stocks, while (Buy*, BA1 to BA3) appear generally persistent, we note the Buy* effects recurrence level to go from frequent to sparse in the

BA4 to BA7 successor events zone. Past this point, with some sparse exceptions, it becomes absent in the BA9 to BA20 area. On the other hand, Sell* events present a generally frequent to highly persistent effect on BA1 to BA16.

4.6.2 Data fitting

Having selected a complete descriptive model for each of our three liquid stocks, the second step of our methodology involves estimating their parameters on a trading day basis. In this section, we use the resulting set of estimated values to perform general data fitting analysis. Our models essentially serving descriptive purposes, we do not consider this chapter in the best data fitting race. Indeed, we consider the basic exponential kernel appropriate for the achievement of our general events relationships analysis task. That being said, it remains obvious that each of our model should deliver a minimum data fitting performance to be considered as serious candidates. Fortunately, we demonstrate that various segments of our descriptive models do not only perform well when compared to the basic Poisson homogeneous model, they also offer satisfying performances in absolute terms.

Before the presentation of our descriptive models performance on a dependant event basis, it is important to recall that on each of the 61 trading days for which the parameters are estimated, the 86 dependant events Hawkes model are fundamentally part of a complete multivariate Hawkes process. Indeed, for each of our three stocks, we actually end up with 61 of these global multivariate Hawkes processes. Consequently, we consider important to mention that for each of these complete processes, the spectral radius of the branching ratios matrix is strictly smaller than one. Indeed, on a daily basis, these values range from 0.65 to 0.77, 0.58 to 0.75 and 0.62 to 0.82 for BMW, ADS and SAP. As explained in Section 4.1, such values suggest the stability of the excitation kernels obtained using our estimation methodology on the data of interest.

Widely used for point processes data fitting analysis, Figure 4.4 presents a quantile-quantile plot (Q-Q plot) for each stock and explained event type pair. Since a perfect fitting would involve inter-event compensators to follow a unit mean and standard deviation exponential distribution, each graphic presents this theoretical distribution as the dotted diagonal line against which our descriptive model inter-event compensators are plotted on a quantile basis. For comparison purposes, we also present the quantiles obtained from homogeneous Poisson processes. For each model, inter-event compensators are computed using the corresponding trading day estimated parameters.

In order to investigate the performance gap between of our descriptive models and the homogeneous Poisson models, we follow Rambaldi et al. (2017) and introduce the *adjusted baseline*. In our context, this measure relates to the proportion of an event arrival process that may be related to its constant baseline component, which correspond to μ^m in equation (4.15). Having estimated expression (4.22) parameters on a daily basis, for each explained event $m \in S$, we refer to $\hat{\mu}_d^m$ as the Hawkes intensity process estimated baseline on trading day $d = 1, \dots, 61$. For the same event type and trading day, $E[\lambda_d^m(t)]$ reports the arrival process unconditional intensity expectation as defined in equation (4.17). Therefore, we define the adjusted baseline as $\hat{\mu}_d^m / E[\lambda_d^m(t)]$, which, as introduced before, reports the proportion of trading day d event m intensity attributable to its constant baseline. Hence, in the absence of Hawkes effect in a descriptive model event arrival processes that would correspond to a Poisson process, this measure would take the constant value one. For each event type, Table 4.3 reports $\overline{\hat{\mu}_d^m}$ and $\overline{\hat{\mu}_d^m / E[\lambda_d^m(t)]}$, the baseline and adjusted baseline daily average. Figure 4.5 complement this information with a visual presentation of most LOB events adjusted baseline daily average.

Beginning with Trade w/ and w/o BPI events, Figure 4.4 Panel I Q-Q plots indicate that despite a performance appearing superior to that of homogeneous Poisson models, this segment of our descriptive models present some weaknesses regarding data fitting. Lu and Abergel (2018) results suggest that these specific events arrival processes could benefit from a double exponential kernel to improve these performance. However, despite these mitigated results, we still consider the simple exponential kernel adequate in our global analysis context. Given our methodology, nothing indicates that the identification of the presence or absence of relationships affecting the arrival of Buy*, Sell*, Buy and Sell events would be affected by the use of a more advanced kernel to describe their intensity processes. Moreover, it is essential to note that despite the fact that a Hawkes process model is considered as a whole, during the estimation step, each events arrival process modeling remains independent of that of the others. Indeed, this representation independence is enforced by the fact that explanatory events occurrence times are exogenous to each dependent event intensity processes. Consequently, in our descriptive context, Trades w/ and w/o BPI imperfect data fitting have no effect on the other 82 LOS and LOC successor event types present in our global framework, even if these former predecessor events appear to sometime play important roles in the latter successor events arrival.

When it comes to limit order submission and cancellation events, Figure 4.4 Q-Q plots presents some interesting trends relative to the data fitting performance of our descriptive models and their Poisson homogenous counterparts. Beginning with LOS w/ BPI, Panel I shows an important performance improvement over Trades. Indeed, despite their previously described particularities, BA1* and AA1* appear adequately represented by our descriptive models. Further into our event types set, Panels II to V show that as we move from the highest to the deepest price levels of occurrence, our models LOS w/o BPI and LOC events arrival processes data fitting performance exhibit some general trends. Despite the facts that these tendencies appear not perfectly monotonic in events depth of occurrence and that each stock presents some idiosyncrasies, by dividing our LOB price levels into three depth segments, it is possible to highlight general patterns that apply to our three liquid stocks. First, our models absolute data fitting performance appears decreasing as we increase in depth level of occurrence. They exhibit their best performance regarding events occurring on low depth levels 1 to 5 with visual results that we consider highly adequate. Beyond this point, our models performance appears to gradually deteriorate. Indeed, Panels II to V show mixed performance regarding events occurring on the order book middle segment consisting in price levels 6 to 11. Despite not catastrophic, events taking place on depth levels 6 to 11 generally exhibit poorer data fitting than what have been observed on lower depth levels. Exceptions are observed for BA10 (AA10) and BC11 (AC11) that present satisfactory results. Afterward, for events taking place on depth levels 12 to 20, our models data fitting capacity appears generally suboptimal. It is interesting to note that in absolute terms, LOS w/o BPI and LOC events fitting performance present very similar characteristics. From an events arrival dynamics point of view, we consider fortunate that disregarding their types [72.2%; 78.8%; 66.8%]

of our identified LOB events take place on depth levels 1 to 5, which correspond to the segment on which our descriptive models present their best performance. On the other hand, we also consider fortunate that only [7.0%; 4.1%; 8.1%] of our identified LOB events appear to occur on depth levels 12 to 20, which correspond to the LOB segment over which our models shows their least interesting data fitting performance.

The second tendency relates to our descriptive models performance relative to Poisson homogeneous models. As shown in Figure 4.4 Q-Q plots the data fitting gap between the two tested frameworks appears to decrease as we increase in LOS w/o BPI and LOC events depth level of occurrence. In addition to being attributable to our selected models deteriorating absolute performance, this trend may be attributed to the improving performance of the Poisson homogeneous models. As before, these tendencies monotonicity being imperfect, they are better perceived when we split the LOB into three depth segments. Back to the low depth levels 1 to 5 events, it is actually possible to note that while our LOS w/o BPI and LOS events descriptive models exhibit their best performance, the Poisson homogeneous models display their poorest results. We relate these visually important gaps to our descriptive models processes low adjusted baseline values. Figure 4.5 shows that among all events represented, the descriptive models intensity processes of those taking place on the five lowest depth levels present the smallest average adjusted baseline with means of [0.28; 0.21; 0.30] for LOS w/o BPI and [0.22; 0.23; 0.32] for LOC events. Such low values indicate a relatively small Poisson homogeneous contribution and large Hawkes effects contribution to their arrival. Indeed, we have to keep in mind that in our current framework, for any event $m \in S$ and trading day $d = 1, \dots, 61$, $1 - \hat{\mu}_d^m / E[\lambda_d^m(t)]$ corresponds to the proportion of event m arrival that may be related the Hawkes effects involved in its arrival process. For low levels LOS w/o BPI and LOC events, not only are these effects numerous, but they seem to bring relevant information to the models, which leads to our interesting data fitting results both relative and absolute. Once again, events taking place on the LOB middle segments present mixed results. Regarding these events which take place on depth levels 6 to 11, the decreasing performance gap relative to Poisson homogeneous processes appears mostly attributable to these last framework data fitting improvement. Figure 4.5 reveals some average adjusted baselines increased values with means of [0.46; 0.48; 0.41] for

LOS w/o BPI and [0.41; 0.49; 0.48] for LOC events. This suggests a tendency for these middle depth events arrival to be slightly more driven by the constant components of our Hawkes processes based descriptive models, which may be related to the reduced visual Q-Q plot gap between them and Poisson homogeneous.

Regarding the events occurring in the deep LOB segment, Figure 4.4 Q-Q plots show that overall, the quantiles related to our descriptive models events arrival processes are close to those of the Poisson homogeneous processes. We explain this global convergence by the reduced number of events relationships involved in the arrival process of the events taking place in this segment. Indeed, as previously presented, these descriptive models processes include a maximum of two Hawkes components. To the extreme, our descriptive models arrival processes for LOS w/o BPI taking place on depth levels 17 to 20 include no Hawkes effects at all, which, as claimed before, leaves us with pure Poisson homogeneous processes. However, unlike events taking place on the two lowest LOB segment, we note an interesting difference between LOS w/o BPI and LOS events in the fact that while quantiles related to the events of the former category almost overlap those of Poisson homogeneous, Panels II to V Q-Q plots show a gap between quantiles related to events of the latter category and those of Poisson homogeneous. Despite subtle, this distinction suggests that on high depth levels, LOC events fitting performance stands out more from Poisson homogeneous than LOS w/o BPI. Leaving the actual events arrival dynamics implications to a subsequent section, we relate this difference to the more important constant baseline participation in LOS w/o BPI than in LOC events arrival processes that is observed in Figure 4.5. Disregarding the processes including no Hawkes component, this figure shows mean adjusted baseline averages of our descriptive models LOS w/o BPI events arrival of [0.75; 0.72; 0.76] while LOC events counterpart values equals to [0.52; 0.46; 0.50], which are closer to the values previously observed for events taking place in the middle depth LOB segments. These numbers suggest a less important constant baseline (more important Hawkes effects) contribution to LOC events arrival processes than to LOS w/o BPI. We consider these elements as a potential explanation for our descriptive models LOC events data fitting performance relative to Poisson that appears to be better than those of LOS w/o BPI when such events take place on depth levels 12 to 20.

4.6.3 Estimated parameters

Having derived the adjusted baselines from the actual estimated constant baseline parameters, we now focus on Hawkes effects related parameters. On a stock basis, we define events $m \in S$ and $n \in \hat{S}^m$, where $\hat{S}^m \subseteq S^m$ corresponds to the set of explanatory events that have been selected for event m descriptive model intensity process. Back to expression (4.22), these (n, m) relationships are daily characterized by the $\alpha_d^{(n,m)}$ and $\beta_d^{(n,m)}$ parameters of which we denote the estimated values as $\hat{\alpha}_d^{(n,m)}$ and $\hat{\beta}_d^{(n,m)}$. Table 4.4 aggregates these parameters by providing their daily average. Each 2-values cell relates to a selected event n exponential Hawkes effect on an event m arrival. For each (n, m) effect, $\overline{\hat{\alpha}_d^{(n,m)}}$ relates to the initial intensity increment and $\overline{\hat{\beta}_d^{(n,m)}}$, to the exponential decay factor. Although not always speaking for themselves, these values represent the core of our multivariate Hawkes events dependency structure.

Regarding the temporal aspect of our several exponential Hawkes effects, Table 4.4 shows that for our liquid stocks, $\overline{\hat{\beta}_d^{(n,m)}}$ ranges from [2.70 to 979; 1.58 to 903; 11 to 946]. These values corresponding to exponential decay factors, it may be difficult to appreciate their implications. Therefore, we focus on the more interpretable half-life period (HL) as Hawkes effects duration measure. However, in order to avoid overweighting outlier $HL_d^{(n,m)}$ values in the (n, m) Hawkes effect duration analysis, instead of computing the actual half-life period daily average $\overline{ln(2)/\hat{\beta}_d^{(n,m)}}$, we use $ln(2)/\overline{\hat{\beta}_d^{(n,m)}}$ in which equation (4.7) is applied on the exponential decay factor daily average.[13] In this context, Figure 4.6 presents these values cumulative distribution functions for the [374; 328; 366] Hawkes effects that have been selected as parts of our descriptive models. With [4.95 ms; 5.91 ms; 5.64 ms] median values and [17.4 ms; 27.5 ms; 14.7 ms] 90th percentiles, these measures suggests important proportions of individual Hawkes effects to remain effective on very small time spans. Although useful as a general duration indicator, we have to keep in mind that effects subject to exponential decay remain effective past their half-life period. As said before, it actually takes about 6.64 half-life periods for these effects to lose 99% of

[13] Outlier $HL_d^{(n,m)}$ values result from irregular $\hat{\beta}_d^{(n,m)}$ estimated parameters that may originate from estimation convergence issues or trading days exceptional idiosyncrasy.

their strength. From the previous median half-life periods, it is easy to establish that 50% of our daily estimated events effects almost totally vanish within [33 ms; 39 ms; 38 ms]. Moreover, it is possible to claim that about [99.4%; 96.2%; 100%] of our estimated Hawkes effects dismiss over time periods shorter than 500 ms, which is considered by Moallemi and Sağlam (2013) as a reasonable estimation of the human reaction time. For technical considerations such as the exponential kernel shape and the mutually-exciting capacity of our Hawkes processes based models, we consider more or less relevant to relate Hawkes effects average effectiveness periods to actual reaction times to events. However, from our point of view, such generalized low average values may suggest algorithm involvement in our three liquid stocks LOB events arrival dynamics.

When it comes to instantaneous increment in events arrival intensity related to Hawkes effects, Table 4.4 shows that $\overline{\hat{\alpha}_d^{(n,m)}}$ values range from [0.31 to 170; 0.52 to 181; 1.18 to 153] events per second with [10.27; 9.49; 8.9] median values. These numbers represent [1.7 to 822; 1.4 to 1368; 5 to 1091] times the involved predecessor event unconditional expected intensity $E[\lambda_m^d(t)]$ with [74; 58; 96] median values[14]. Like exponential decay factors and effects HL period, although suggesting potentially important instantaneous intensity increases, it is highly difficult to interpret the $\overline{\hat{\alpha}_d^{(n,m)}}$ values by themselves. Indeed, to adequately quantify an actual predecessor event n impact on a successor event m arrival on a trading day d, the initial intensity surge has to be placed in its temporal context, which relate to the exponential decay factor. Back to our daily aggregated context, Figure 4.7 Panel I shows a relation in the magnitudes of $\overline{\hat{\alpha}_d^{(n,m)}}$ and $\overline{\hat{\beta}_d^{(n,m)}}$, which makes comparing Hawkes effects with different effective periods very difficult. Adding the fact that several hundred relationships have been identified, it becomes obvious that an analysis directly based on these two parameters for each effect would be inconvenient. In this context, the branching ratio (BR) becomes an interesting measure of effect strength. As defined before, $BR^{(n,m)}$ corresponds to the effect of a single event n occurrence on event m compensator. Therefore, it represents the expected number of event m that may

[14] We exclude (BA1*, BC11) and (AA1*, AA11) relationships that we identify as outliers with $\overline{\hat{\alpha}_d^{(BA1*,BC11)}}$ values of [362; 322; 324] and $\overline{\hat{\alpha}_d^{(AA1*,AC11)}}$ values of [353; 317; 323].

be related to the arrival of an event n. Back to our daily aggregated framework, Figure 2.7 Panel II shows that unlike what we have observed for $\overline{\hat{\alpha}_d^{(n,m)}}$, the magnitude of $\overline{BR_d^{(n,m)}} = \overline{\hat{\alpha}_d^{(n,m)}/\hat{\beta}_d^{(n,m)}}$ appears stable in $\overline{\hat{\beta}_d^{(n,m)}}$, which makes our several Hawkes effects comparable when using this metric. Additionally, $BR_d^{(n,m)}$ s.t. $d = 1, \dots, 61$ appears more stable than $\hat{\alpha}_d^{(n,m)}$ and $\hat{\beta}_d^{(n,m)}$. Figure 4.8 provides visual general examples for these metrics dispersion with respect to their daily average. Panel I presents [0.29; 0.28; 0.29] and [0.32; 0.28; 0.27] coefficients of variation for $\hat{\alpha}_d^{(Buy*,BA1*)}$ and $\hat{\beta}_d^{(Buy*,BA1*)}$. On the other hand, the $BR_d^{(Buy*,BA1*)}$ coefficients of variation are [0.16; 0.16; 0.15], which is [1.8; 1.7; 1.9] and [1.9; 1.7; 1.8] times smaller than those observed for the previous two estimated parameters.[15] Such information suggests $\overline{BR_d^{(n,m)}}$ to be more representative of the daily results than $\overline{\hat{\alpha}_d^{(n,m)}}$ and $\overline{\hat{\beta}_d^{(n,m)}}$ averages, which is fortunate for our next section analysis.

4.7 Events dynamics

In this section, we use our descriptive models estimation results to analyse different aspects of the events arrival dynamics. We also attempt to identify some patterns that may be related to trading strategies and market participants behaviors. In our exponential Hawkes framework, different events relationships patterns are possible. Indeed, it is possible for multiple predecessor events of the same type to combine and increase the arrival rate of a successor event as long as it is possible for this task to be performed by precursor events of different types. We first have the self-exciting case in which, since $n = m$, the (n, m) effect leads to events of the same type clustering phenomena. We also observe the second situation where $n \neq m$ and the presence of the reciprocal (n, m) and (m, n) relationships potentially translate into mutually-exciting situations. We note a third situation in which $n \neq m$ but only one of the (n, m) or (m, n) relationships is present. And finally, we observe several situations that we qualify as plurally-exciting in which

[15] Figure 2.8 Panel II presents the mirror (Sell*, AA1*) relationship example where the coefficients of variation for $\hat{\alpha}_d^{(\mathrm{Sell}*,\mathrm{AA1}*)}$ and $\hat{\beta}_d^{(\mathrm{Sell}*,\mathrm{AA1}*)}$ are [0.28; 0.26; 0.30] and [0.30; 0.26; 0.31], which are[1.9; 1.7; 2.0] and [2.0; 1.8; 2.0] times smaller than those [0.15; 0.15; 0.15] $BR_d^{(\mathrm{Sell}*,\mathrm{AA1}*)}$ values.

more than two events are involved in an arrival dependency structure. Taking the simple example of the simple self-exciting case, the same alpha and beta parameters are used to characterize the timing and strength impact of an orphan event occurring in an inter-exciting period as well as for the third event of a cluster. The same applies to another example in which three events are defined: $m \in S$ and $n_1, n_2 \in S^{m*}$ such that $S^{m*} \subseteq S^m$. As before, this definition involves the presence of (n_1, m) and (n_2, m) relationships, which we define as plurally-exciting. In this case, we denote essentially five scenarios through which events n_1 and n_2 may have an impact on event m arrival. We have the orphan n_1 or n_2 event, multiple n_1 or n_2 occurrences and finally, the cases where we observe a combination of n_1 and n_2 events. Since each of theses events arrival structures may origin from various trading strategies, on a daily basis, the maximum likelihood estimation methodology have to result in the best overall data fitting. Therefore, while some relationships estimated parameters suggest the successor event arrival rate to reach very high levels over small periods, we have to keep in mind that they may be generally part of intensive excitation periods involving several types of predecessor events. Therefore, with some exceptions covered in the next section, we consider generally irrelevant to focus on a predecessor event single occurrence effect on a successor event. Indeed, since dependant on the arrival context, from our point of view, isolating the marginal effect of an event n occurrence on event m arrival probability requires too many strong assumptions such as the certainty that no other event affecting the arrival of event m would occur over an arbitrary chosen time period to provide reliable results. In this context, since constant through all possible states, once again, the branching ratio (BR) represent a choice measure for our analysis of the events arrival dynamics.

However, while providing information on a single event n occurrence effect on event m expected number, this ratio provides no information on the total events n involvement in events m arrival over a complete trading day. Therefore, comparing event relationships using BR as the only measure may lead to some misinterpretations. Indeed, for the same $BR_{(n,m)}$ value, the (n, m) relationship implications differ given that event m is less, equally, or more frequent than event n. In this context, to complement the branching ratio in our relationships comparison, we use the following expression to define the adjusted branching ratio, which is based on Rambaldi et al. (2017) adjusted kernel norm :

(4.23)	$$\widetilde{BR}_{(n,m)} = BR_{(n,m)} \frac{E[\lambda_n(t)]}{E[\lambda_m(t)]}.$$

Under this expression, $BR_{(n,m)}$ is *adjusted* through its multiplication by the ratio of event n on event m unconditional expected arrival intensities. The resulting adjusted branching ratio $\widetilde{BR}_{(n,m)}$ reveals the proportion of event m intensity that may be related to occurrences of event type n. In opposition to the previously defined *adjusted baseline* who provides the proportion of event m arrival that may be related to its constant baseline μ_m, the sum of its effects *adjusted branching ratios* provides the share that may be related to its Hawkes components.

From this point, we mainly focus on the effects that have been considered recurrent enough to be included in our descriptive models. Since they have met our selection criteria at least on every other trading day, we believe these relationships to act as potential pieces of the events arrival dependency structure, which is summarized in Figure 4.9. Adopting a top-down approach, we begin with Panel I that presents a very general picture of the interrelations among our high level events categories. Afterward, in order to generalize our observations, we work on the basis of relationships classes. As presented in Panel II, these classes group relationships by involved successor and predecessor event types and affected book sides. Depending on the involved event categories, such relationship classes may encompass one events effect or more. We assign them an identifier that goes from A1 to S1. Based on the LOB side affected by the involved events, each class is present in two variations. As an example, we consider that Bid LOS effects on Bid LOS and Ask LOS effects on Ask LOS events belong to the dual relationship classes identified P1, which is present twice in Panel II. Once again, because of the strong symmetry observed so far in the relationships affecting the bid and ask LOB sides, we attempt to simplify the presentation by essentially focussing on the buyers related events. Therefore, while Panel II presents both variations of our relationship classes, a single version is visually represented in Panels III and IV. Panel III exposes the dynamics related to Buy*, Buy and BA1* events. Panel IV performs the same task for Bid LOS w/o BPI and Bid LOC events. Supplementing Figure 4.9 as the core of the next subsections, for our liquid stocks, Table 4.5 presents the actual average branching ratios (BR) and adjusted branching ratios ($\widetilde{BR}$) characterizing the individual Hawkes effects relating Trades, LOS w/ BPI and, LOS and LOC events taking place up to depth level 11.

4.7.1 Trades

As pointed out before, our descriptive models suggest Trades and Trades w/ BPI events arrival to be affected by a limited number of events. Although few in number, these effects identified as recurrent by our methodology bring their share of LOB dynamics elements. We relate (Buy, Buy*) (A1) and (Buy, Buy) (C1) relationships to the trades clustering phenomena covered by Hewlett (2006). The self-exciting component C1 suggests the probability of observing a Buy event to be increased after the occurrence of an event of the same type. Similarly, C1 indicates that following a Buy event occurrence, the probability of Buy* event is also increased. These Hawkes effects being additive, more Buy events taking place within their effective time period lead to even higher Buy and Buy* occurrence probabilities. Despite the fact that they do not take the number of involved shares in each situations into account, A1 and C1 are consistent with a trading strategy consisting into splitting the execution of a large parent order into small child orders. We use the simple example of an investor intending to acquire 1000 shares who chooses to perform this task by submitting five consecutive 200 shares child aggressive orders instead of a single 1000 shares order. In a first scenario where the best ask price level contains more than 1000 shares, this execution strategies would result in 5 consecutive Buy events, which is consistent with the self-exciting C1 relationship. In a second scenario where the best ask price level contains exactly 1000 shares, four Buy events would be first observed, followed by a Buy* event. In addition to C1, this case would also involve the A1 relation that suggests the short term probability of observing a Buy* event to be increased after the occurrence of a Buy event. Finally, the same effects combination would be involved in a third situation where the best ask price level would contain less than 1000 shares, which would lead to one or more of the child aggressive orders to execute against liquidity available on at least one higher depth level. In this context, the Buy* event would be potentially followed by one or more Buy events that would be part of the whole cluster. This first events arrival dynamics case illustrates the complexity involved in relating our descriptive models results to actual trading situations because of the limited number of parameters available to characterize several complex situations.

As presented in Figure 4.9, Trades arrival intensity also appear affected by LOS w/ BPI events. Detailed in Panels II and III, (AA1*, Buy*) (B2) and (AA1*, Buy) (D1) suggest an increase in the expected number of Trade with and without BPI after the submission of an ask limit order inside the spread creating a new best price ask depth level. On the other hand, (BA1*, Buy*) (B1) suggests a similar phenomena to follow the arrival of a bid limit order inside the spread. Unlike previous A1 and C1 relationships, our liquid stocks descriptive models present some discrepancies when it comes to the effects of LOS w/ BPI. While B2 have appeared persistent enough to be included in our models for each of our liquid stocks, B1 have not met our selection criteria for SAP, with the consequence of not making its descriptive model. The same applies to D1 that is not part of the ADS model. Nevertheless, we examine the interpretation of these last two effects insofar as, in addition to having satisfied our selection criteria for two of the three stocks, back to Table 4.2, we do not observe an absolute absence of relationship regarding the third one. With the most important BR values among these relationships categories, we relate the persistent B2 relationship to situations where the submission of a best price improving ask limit order (AA1*) would be followed by its rapid total execution (Buy*). In the same way, we relate D1 to the eventual partial consumption of a similar incoming limit order. When considered with C1 and A1, it is possible to relate D1 and even B2, to a Buy events cluster that may eventually lead to the total consumption of the new price level. Finally, we relate B1 to a different dynamics. In this case, we consider the mid-quote price increase related to a new best bid depth level as an incentive to acquire the liquidity available at the best ask price quickly. From Table 4.5 Panel I, it is interesting to note that despite smaller BR values, the $\widetilde{BR}$ values ranges for B1, B2 and D1 are comparable to those of A1 and C1, which suggests the contribution of these Hawkes effects in the Trades arrival processes to be similar. From equation (4.23), it is possible to relate this observation to the fact that BA1* and AA1* are more frequent than Buy* and Buy, which appear to compensate for their more limited individual occurrence effects.

4.7.2 Limit orders submissions with best price impact

Panels II and III of Figure 4.9 shows that LOS w/ BPI arrival rate may be affected by the occurrence of events from our four main categories. From a mechanical point of view, it is not surprising to note that, as indicated by their BR values in Panel II of Table 4.5, Trades w/ BPI appear to have the more important events occurrence individual effects. These events resulting in an increased bid-ask spread, more room becomes available for the creation of a new best price level on both side of the book. Despite the fact that our models do not explicitly take the magnitude of the spread growth into account, we know for sure that it corresponds to at least one tick. First, we consider the highly persistent (see Table 4.2) (Buy*, BA1*) (E1) relationship as a part of a price following dynamics in which the mid-quote price increase caused by the Buy* event may be followed by the creation of a new best bid price level. On the other hand, we consider (Sell*, BA1*) (E2) as part of a LOB resilience dynamics where the total execution of the previous best bid price level resulting in a Sell* event may be followed by the creation of a new one, through a limit order submission that would be identified as a BA1* event.

When it comes to Trades w/o BPI effect on BA1* (F1), we consider this relationship in a mutually-exciting context. Indeed, because of the Hawkes effects additive nature and the previously described clustering phenomena related to Buy event that may lead to a Buy*, we consider possible for Buy events to be part of potential dynamics leading to the formation of a BA1* event favorable environment, once again, in a price following context. Despite less important in terms of individual events occurrence effect, G1 to I1 indicates that some LOS and LOC events taking place around the best price levels may have effects on BA1* arrival. First, (BA1*, BA1*) (G1) suggests the possibility for BA1* events to occur in clusters. We consider successive BA1* events consistent with a book liquidity replenishment situation that may follow, as an example, a Sell* event (E2) resulting from a transaction having consumed more than one bid price levels. Regarding (BA1, BA1*) (H1), we consider the presence of its reciprocal (BA1*, BA1) (N1) relationship (see Table 4.5 Panel II) as an indicator that BA1 and BA1* events occurrence may be favorited under similar circumstances. Both event types involve added liquidity on the best bid price level, actual or new. Finally, we relate I1 and I2 to the bid-ask spread enlargement that may result from BC1 and AC1 events in situations where the cancelled limit order is the only constituent of the best price level. Shown in Table 4.5 Panel II, we consider the total absence of (BA1*, BC1) and (AA1*, AC1) relationship for our liquid stocks as an indication that, in general, the cancellation of limit order creating a new best bid or ask price level does not immediately follow its submission.

4.7.3 LOB events

Having highlighted some interesting characteristics of Trades and LOS w/ BPI events arrival process, the main goal of this section is to perform a similar analysis for LOB events. Depending on the predecessor events, we observe relationships involving LOS w/o BPI and LOC as the successor event to present different characteristics. First, we study the effects of Trades and LOS w/ BPI on LO submissions and cancellations. Afterward, we achieve a similar task through the analyze of LOB events effects on each other arrivals. Finally, we take a look at each of these event categories contribution to the LOB events arrival.

Figure 4.10 presents the effects of Trades w/ BPI, Trades w/o BPI, and LOS w/ BPI on LOS w/o BPI and LOC events taking place on depth levels 1 to 20 in terms of branching ratios. In line with our previous visual representations, J1 to K2 relationship classes are detailed through the inner graphics displaying the corresponding labels. Still working from the buyer point of view, we focus on Panels I, III and V that exhibit the effects of a single occurrence of Buy*, Buy and BA1* events. By concentrating on these predecessor events similarities, differences, and effects on LOB events arrival, we attempt to characterize the dynamics surrounding their occurrence. Basically, we relate the Buy*, Buy and BA1* events through the fact that they all result from an aggressive buyer action. Consuming at least one complete price level, we consider a Buy* event as very aggressive. Similarly, resulting from the partial execution of the best price level, we qualify the Buy event as aggressive. And, although it does not immediately lead to an actual transaction, since it results from the submission of the most aggressive limit order at a given time, we consider BA1* to immediately follow the two former events on our aggressiveness scale.

When not related to only one of these events, the other properties relevant to our dynamics description are shared on an events pair basis. First, while Buy* and Buy affect the ask side of the book, BA1* have an impact on the bid side. Second, Buy* increases, BA1* decreases, and Buy has no effect on the bid-ask spread. Third, unlike Buy that does not affect the price structure, by respectively leading to a change in the best ask and bid prices, Buy* and BA1* increase the mid-quote price. When it comes to the effects on LOB events arrival, we identify these characteristics as more important than the actual liquidity execution involved in Buy* and Buy. Figure 4.10 suggests BA1* effects on LOB events arrival to be more important in number and in intensity than those related to Buy. However, one may expect Buy event effects to be closer to these of Buy* since they both involve actual liquidity execution, which appears inaccurate according to our descriptive models results. Indeed, while Buy* stands out as the most important predecessor event in terms of individual occurrence effects, it is immediately followed by BA1*. Since Buy is left behind, we identify the mid-quote price increase and the change in the LOB structure specific to the other two events as important determinants of the events arrival dynamics. With these LOB impacts in mind, we first distinguish two sets of similar relationship classes involving our three predecessor events. While J1, L1, and N1 suggest Buy*, Buy, and BA1* to affect bid LOS arrival, K2, M1, and O2 do the same regarding ask LOC.

Beginning with bid LOS, as shown in Figure 4.10, the most important effects in terms of number of affected events and branching ratios belong to J1 and N1. These relationships suggest Buy* and BA1* events to be immediately followed by increased probabilities of bid LO submissions respectively taking place on depth levels 1 to 5 and 1 to 7. We relate these effects to eventual buyers submitting their limit orders in an environment where the stock price may appear to be increasing. Indeed, Buy* and BA1* events being accompanied by a mid-quote price increase, we consider these potential buyers to submit their limit orders behind this new price, in an attempt to follow this possible upward trend. At this point and all over the current analysis, we have to keep in mind that we denote increases in the expected number of bid LOS events following Buy* and BA1* occurrences, which highly differs from an actual systematic limit orders flow simultaneously taking place on these multiple depth levels. We assume that the depth levels on which such potential limit orders would be submitted would be determined by their owners patience level and trading strategy. Regarding the Buy predecessor event, as presented in Figure 4.10 Panel III, L1 suggests it only affects bid LOS arrival taking place on depth levels 1 to 3 through effects that present smaller BR values than those attached to Buy* and BA1*. We relate these less important effects to the previously described absence of individual Buy event structural LOB effect. The transactions leading to the identification of these events only partially affect the best ask price level. However, we relate their presence to the fact that, as for Buy*, Buy events occurrence may suggest the presence of an impatient buyer, which may sometime be interpreted by patient sellers as a potentially increasing price signal. Also, we have to consider the previously established self-exciting nature of the Buy event that could lead to an addition of these effects over short time periods.

Regarding the second group of relationships similar for our three predecessor events, as claimed before, K2, M1 and O2 suggest Buy*, Buy and BA1* to have an effect on ask LOC events arrival on various depth levels. These events potentially carrying signs of an increasing price, we consider the fact that they appear to favor the cancellations of sell limit orders intriguing. Indeed, it suggests that some market participants tend to cancel their limit orders as their execution probability increases. From our point of view, this phenomenon may be related to different situations. First, we consider possible for some orders to be owned by market makers with the intention of providing liquidity in the LOB without a real interest in seeing their orders executed. Second, they may be part of complex trading strategies potentially involving LO submissions and cancellations on both sides of the book. Finally, Buy*, Buy and BA1* suggesting the presence of impatient buyers, we consider possible for patient sellers already present on the ask side to cancel their limit order in an attempt to take more advantage of the potentially increasing price. In this last case, the cancelled limit orders would be eventually resubmitted with higher prices. Although Figure 4.10 suggests these relationships to be related to Buy*, Buy and BA1* events, it indicates Buy* to have the most important underlying effects. Indeed, while appearing to affect ask LOC events up to the ninth depth level, Buy* effects seem atypically strong for LOC taking place on levels 1 to 3. Actually, with average branching ratios generally above 0.8, the highest BR values observed in our entire system belong to Buy* effects on AC1 and AC2 events arrival. We relate this important K2 relationship class to the fact that the Buy* event directly affect the ask side of the book by completely consuming one or more price levels, which is not the case for Buy and BA1*. A Buy* event occurrence has for consequence that all existing ask limit orders rank relative to the best ask price is improved by the number of price levels actually consumed by the underlying transaction. In our framework context, this rank correspond to the depth level number. No matter the trading strategies behind K2 effects, it is a certainty that the potentially involved limit orders have become closer to the best ask price before their cancellation. In this context, we consider possible for some systematic trading strategies to use limit orders depth level as a trigger to launch their cancellations.

As presented through J2, Buy* events favor ask LOS through numerous relationships which, in some cases, may also be qualified as important in BR values terms. While observed up to depth level 16, these effects present more important BR values for AA1 to AA10. Back to Figure 4.10 Panels III and IV, we observe a complete absence of similar effects involving Buy and BA1*. In line with a possible upward price trend, we relate these increased probabilities to eventual patient sellers submitting limit orders in an attempt to obtain an even better price than the actual new best ask price. The different depth levels over which the LOS probabilities are affected suggest various levels of patience, which may depend on the concerned investors characteristics and their trading strategies. Considering the previous scenario where a still patient seller would attempt to take a chance at a possibly increasing price. In an ask side reorganization context, a cancelled ask limit order could be submitted on a higher depth level. Despite speculative because of the lack of information sequentially relating K2 and J2, a part of J2 induced probability increases could be attributed to these resubmitted orders.

We end the analysis of Trades and LOS w/ BPI impacts on LOB events arrival with relationship classes K1 and O1, which corresponds to Buy* and BA1* effects on bid LOC. Panel III shows such effects to be totally absent when it comes to the Buy predecessor event. While K1 and O1 involve bid LOC probabilities increases up to depth level 19, we relate the absence of such effect regarding Buy to the fact that by definition, it does not change the best bid and ask prices and by extension, the mid-quote price. We relate bid LOC events that may arise in the K1 and O1 contexts to situations where limit orders would be cancelled by buyers observing an adverse move in the stock price. The unusually high number of depth levels over which LOC probabilities appear affected suggests various patience levels among the potentially concerned investors over time. Regarding K1, similar to what happens on the ask side where existing orders become closer to the best ask price as represented by K2, a Buy* event has for consequence that the limit orders already present on the bid side become farther from the best ask price. Therefore, depending on an investor patience level and trading strategy, a given distance from the best ask or mid-quote price may act as a threshold to trigger a limit order cancellation. From our point of view, unlike K2 where the cancelled limit orders were getting too close to the action for still patient sellers, K1 cancelled limit orders seem to get too far from the action to be kept alive by patient buyers seeing their execution probability reduced. In the O1 case, we observe some of the effects to be characterized by BR larger than those seen in K1. We explain this situation by the fact that unlike Buy* which affects the ask side of the book, BA1* events have a direct impact on the bid side through the addition of a new depth level that becomes the best price level. All existing bid limit orders are then affected by being automatically shifted one price level away from the best bid price. These effects strength suggests a limit order distance from its book side best price to be more monitored and used as a cancellation trigger than its distance from the opposite book side best price. Moreover, we consider the (BA1*, BC11) important BR value as an indication that the 11th price level relative to the best price may be used by several algorithms as a threshold to determine that a limit order has become too far away (above level 10) from the action and, must be cancelled. We relate this strong relationship to the particularly good data fitting performance of our descriptive models regarding BC11 and AC11 (see Figure 4.4), which suggests an adequate representation of this dynamics. Additionally, O1 graphic suggests that reaching depth levels 2, 4 and 6 could also be single out as limit orders cancellation triggers.

Regarding LOB events effects on each other arrival, we summarize these impacts through relationship classes P1 to S1. Figure 4.10 Panel IV provides a general idea of how bid LOS and LOC are related to other events of the same classes taking place on both sides of the book. In the previously described cases of Trades and LOS w/ BPI events, the absence of reciprocal counterparts to relationship classes J1 to O2 allowed the assumption of some forms of causality. While Trades and LOS w/ BPI appear to affect the arrival of LOB events in different ways, only few relationships going the other way have been observed. Indeed, with the exceptions of H1 to I2 that suggest best price levels LOS w/o BPI and LOC effects on BA1* and AA1* arrival, Trades and LOS w/ BPI generally appear unaffected by LOB events. However, it is possible to observe that some of the P1 to S1 relationship classes, which involve LOB events only, act as each other reciprocal. The resulting mutual-excitation phenomenon affect various events sets going from pairs of individual events of the same type to events groups of different types taking place on both sides of the book. By grouping extracts from Table 4.5, Table 4.6 summarizes the two most extended cases of mutually exciting zones present in our events dependency maps. While Panel I presents interrelations among bid LOS and ask LOC, Panel II does the same regarding ask LOS and bid LOC. Insofar as this phenomenon appears concentrated on the first three price levels, we solely focus on this segment of the LOB. Panel I shows that for each of our three liquid stocks, multiple relationships are present in the (BA1 to BA3, AC1 to AC3) zone, which correspond to the ask Q2 relationship class. In a reciprocal way, the bid R2 class reveals several effects in the (AC1 to AC3, BA1 to BA3) zone. The situation increases in complexity when we also consider the bid P1 and ask S1 relationship classes that correspond to (BA1 to BA3, BA1 to BA3) and (AC1 to AC3, AC1 to AC3). Although the involved individual effects do not present particularly high BR values, we conclude in an important level of interrelation among them. While Table 4.6 examples generally exhibit some individual relationships absence, Panel II shows total interconnection among AA1 to AA3 and BC1 to BC3 events in the BMW case. In this specific situation, all events occurrences seem to affect their own event type arrival and the arrival of each of the other event types present in the set. Although not all involving total interconnection, each Table 4.6 example involves a complex multi-event mutually excitation chains whose detailed interpretation may become very challenging. Consequently, in such situations, it becomes difficult to go further than assuming that the involved events tend to occur under the same circumstances. It is important to note that these circumstances may involve the occurrence of Trades and LOS w/ BPI events, or not. Indeed, although possible that J1 to O2 capture an important part of the Trades and LOS w/ BPI effects on LOB events, in our exponential Hawkes models context, P1 to S1 rely on the same parameters set to describe dependency among LOB events arrival whether or not they are part of a dynamic related to J1 to O2. However, Trades and LOS w/ BPI events being generally less frequent than LOB events, we consider possible for P1 to S1 relationship classes to be more related to situations taking place outside these events dynamics.

As detailed through Panels III to VI, relationship classes P1 to S2 sometime have a *black box* appearance. However, although impossible for us to interpret each individual effect, we are still able to observe some general patterns. First, reminding the complete absence of (Buy, Sell*), (Sell, Buy*), (Buy, Sell) and (Sell, Buy) relationships, we denote the almost total nonappearance of recurrent relationships between bid and ask LOS events. A similar situation is observed regarding bid and ask LOC events, which occurrences also appear unrelated. It is true that we have observed the pairs of relationships (J1, J2), (K1, K2) and (O1, O2) that suggest some of these events to tend to follow the same Trades and LOS w/ BPI events. However, we consider that the difference in the BR values patterns specific to these relationships pairs consistent with this absence of interrelation among the affected events. As claimed before, these absences suggest opposite sides LOS and LOC to occur generally under distinct circumstances. Since these two class of events suggest different strategies and reading of the market conditions, it would be consistent for market participants to do not tend to perform the underlying actions concurrently. At the opposite, P1 suggests that LOS taking place on the same side of the book tend to occur in similar contexts. As exposed by S1, the same phenomenon appears to apply to same side LOC events. As shown in Panels III and V, for both P1 and S1, the intensive mutual-excitation zone previously described appears to take place up to depth level four. Beyond this point, the same side (LOS, LOS) and (LOC, LOC) recurrent effects appear concentrated around the diagonal, which suggest that individual effects mostly relate events taking place on nearby depth levels. While same side (LOC, LOC) relationships appear diffuse and tend to vanish around depth levels 6 to 8, we observe recurrent (LOS, LOS) effects to follow a straight line up to depth level 10. It is interesting to note that this line is different from the actual diagonal that would encompass self-exciting Hawkes effects through which LOS events taking place on a given level would tend to follow each other. In fact, we observe an offset in the (BA3 to BA9, BA4 to BA10) line that suggests LOS occurrences on a given depth level to be followed by an increase in the arrival probability of LOS taking place one level higher. The depth level number being relative, it is unfortunately impossible for us to establish whether there is a price difference between the orders involved in these relationships or not.

Regarding relationships involving limit orders submission and cancellation events taking place on the same book side, since they respectively encompass (LOS, LOC) effects and their reciprocal counterparts (LOC, LOS), for a given book side, we consider Q1 and R1 relationship classes as paired. Back to Table 4.5, Panels III to VI show that R1 includes more recurrent effects than Q1, which suggests a tendency for LOS events to follow LOC. This is consistent with our previous hypothesis suggesting that certain systematic strategies may reposition limit orders following certain changes in market conditions that may be driven by Trades w/ BPI and LOS w/ BPI. Finally, as introduced before, Q2 and R2 relate LOS and LOC events taking place on both sides of the LOB. As presented in Panels III to VI, each of these relationship classes involves an important mutual excitation zone that generally extends up to the third or fourth depth levels. Back to Figure 4.10, we have seen that relationship class pairs (J1,K2), (L1, M1) and (N1, O2) suggest a tendency for opposite sides LOS and LOC events to follow the same Trades and LOS w/ BPI events. The pairwise BR values patterns presenting similarities, we consider these Trades and LOS w/ BPI events arrival as circumstances under which these LOS and LOC events may tend to occur together. Taking place inside these contexts or not, these aggregated reciprocal relationships remain difficult to interpret since as usual, it is impossible to establish if the LOS and LOC events origin from the same market participants or not. However, they would be consistent with any strategy involving the cancellation of limit orders present on one side of the book and submission of new ones on the other side. As for all dynamics discussed in this section, the exact motivation and timing of such actions would depend on their trading strategies.

Having focused on specific types of events relationships, we close this section with an overview of global event categories contribution to the arrival of LOB events. Figure 4.11 shows the average proportions of Trades w/ BPI, Trades w/o BPI, LOS w/ BPI and LOB events to the arrival of LOS w/o BPI and LOC events. For a given category, this measure sums the adjusted branching ratio of the included effects and reports the daily average. No distinction is made between the effects involving predecessor events impacting the LOB bid side and those affecting the ask side. These results are related to those presented in Figure 4.5 as for each successor event and trading day, the effects contribution sum essentially equals to one minus the adjusted baseline value. Proceeding in the usual order, we first observe a general steadiness regarding Trades w/ BPI contribution to the arrival of both LOS w/o BPI and LOC events taking place on depth levels 1 to 8. Beyond this point, we note a general increase in this contribution that afterward, remains generally constant until it vanishes among the highest considered levels. It is interesting to note that despite some very important individual effects, the contribution of Trades w/o BPI appears relatively small. Buy* effects on AC1 and AC2 represent good examples of this situation. Back to Figure 4.10 Panel I, we have seen that they present the most important individual events occurrence impact with average branching ratio values of [0.81 and 0.85; 0.83 and 0.92; 0.74 and 0.79]. However, when it comes to contribution, the average adjusted branching ratios fall to [0.11 and 0.10; 0.11 and 0.10; 0.11 and 0.12]. Using expression (2.23), it is possible to relate this phenomenon the relative infrequence of the Buy* event with regard to BC1 and AC1. This applies to Trade w/ BPI effects on LOS w/o BPI and LOC events taking place on several depth levels. In fact, because of their generally significant individual effects on LOB events and small relative number of occurrences, we identify Trade w/ BPI effects as periodical LOB activity boosters. On the other hand, in line with their small individual impacts, we denote a marginal Trades w/o BPI contribution to LOB events arrival.

When it comes to the LOS w/ BPI contributions to LOS w/o BPI and LOC events arrival, although appearing more important for cancellations, they remain generally similar for both categories regarding events occurring on depth levels 1 to 5. Beyond this point, while rapidly vanishing for LOS w/o BPI, for LOC they show increasing trends as the distance from the best prices also increases. From our point of view, these phenomena suggest that even on the deepest depth levels, orders cancellations may be more related to automatic actions than orders submissions. The important LOS w/ BPI events contribution to these LOC events may once again be related to systematic trading strategies monitoring their own limit orders rank relative to the same or opposite book side best price. And, based on their rank, take the decision to cancel limit orders or not. A good representation of this mechanics could be related to the fact that as seen before, our descriptive models seem to deliver a better fitting performance for LOC than LOS w/o BPI events, especially on the deep price levels.

Back to Figure 4.11, we note a similarly important aggregated LOB events contribution to the arrival of both LOS w/o BPI and LOC events taking place on depth levels 1 to 5. Despite the previously presented small individual (LOB events, LOS w/o BPI) and (LOB events, LOC) relationships contribution, since these effects are numerous to be part of our descriptive models, once combined, they represent the main drivers for this segment of the dynamics. Beyond depth level 5, their cumulated contribution to LOS w/o BPI and LOC arrival becomes different. Indeed, while remaining important determinants of LOS w/o BPI arrival at least up to depth level 10, their contribution to LOC arrival decreases until it vanishes around depth level 8.

4.8 Summary

In this chapter, we analyze the limit order book events arrival dependency structure. Obtained using the methodology developed in Chapter 2, our set of potentially interconnected events consists in Trades, Limit Order submissions and Limit Order cancellations taking place on the first twenty depth levels of the book.

At every step of our process, we observe a generalized symmetry regarding events affecting the bid and the ask sides in terms of Hawkes effects recurrence and strength. We consider this as a sign that similar trading strategies may operate on both sides of the book. Similarly, despite some idiosyncrasies among our three stocks, we note that the absent and the highly recurrent relationships generally turn out to be the same for each of them.

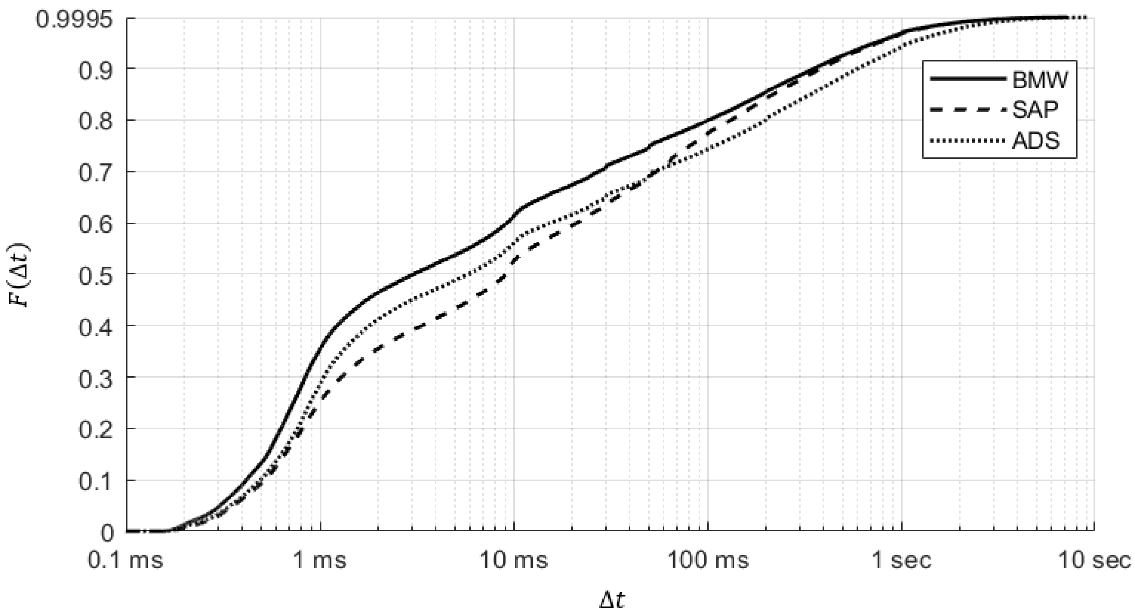
Once our descriptive models estimated, we find that beyond depth levels 1 to 6 where our descriptive models offer a satisfactory data fitting performance for order submissions and cancellations, this performance shows a general decreasing trend as we move in events depth level of events occurrence. On the same depth levels, the Poisson homogenous comparative models show an improving fitting performance. We relate this performance convergence to a decrease in the Hawkes effects contribution to the concerned events arrival that comes together with an increase in the constant baseline contribution.

Through the investigation of the global events arrival dynamics, we confirm that while trades and limit order submissions having an effect on the best price sometimes affect limit order submissions and cancellations up to the deepest segment of the order book, these events appear almost totally unaffected by events taking place beyond the second depth level. We show that while both trades and limit order submissions with best price impact may affect limit order book submissions and cancellation up to the highest depth levels, the effect of their counterpart events without best price impact is less extended.

We also find that both in terms of individual occurrence effects and global contribution, limit order cancellations appear particularly affected by the events having a best price impact. We relate this observation to the possibility for cancellations to be systematically launched on the basis of criteria that may be well represented by our models.

Finally, we observe segments of our events dependency structure over which several event types appear to affect each other’s arrival, which we qualify as mutually-exciting zones. Based on the involved events type and the relationships characteristics, in these cases, we rule out the causality and consider these events as potentially related to the same factors.

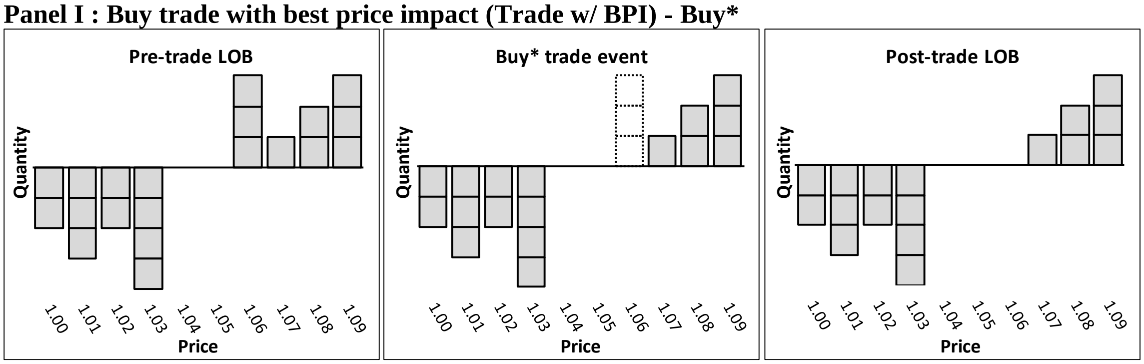
Figure 4.1 LOB inter-updates durations empirical CDF



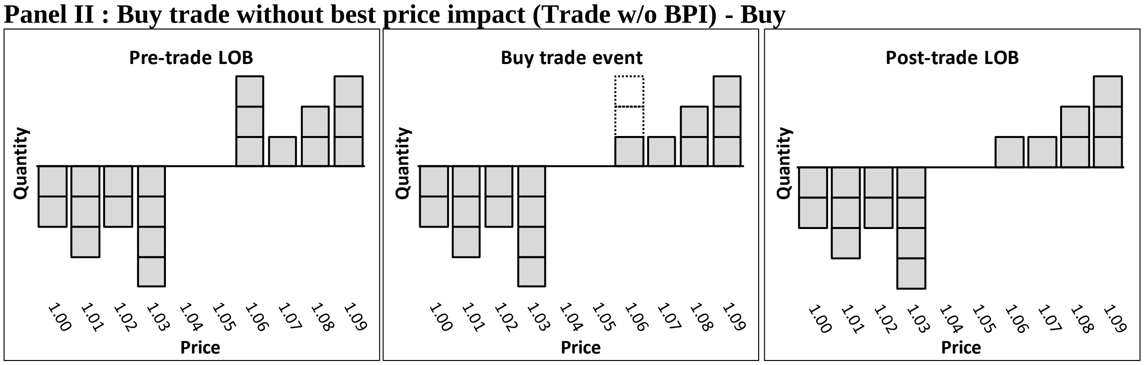
This figure presents the cumulative distribution function of the time between two limit order updates for the BMW, SAP and ADS stocks during the period going from February 1 to April 30, 2013.

Figure 4.2 Fictional events examples and their LOB effects

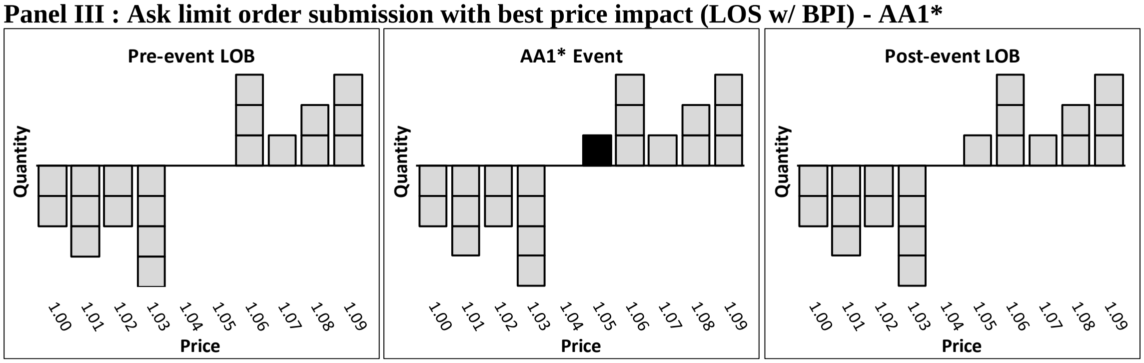
Panel I : Buy trade with best price impact (Trade w/ BPI) - Buy*



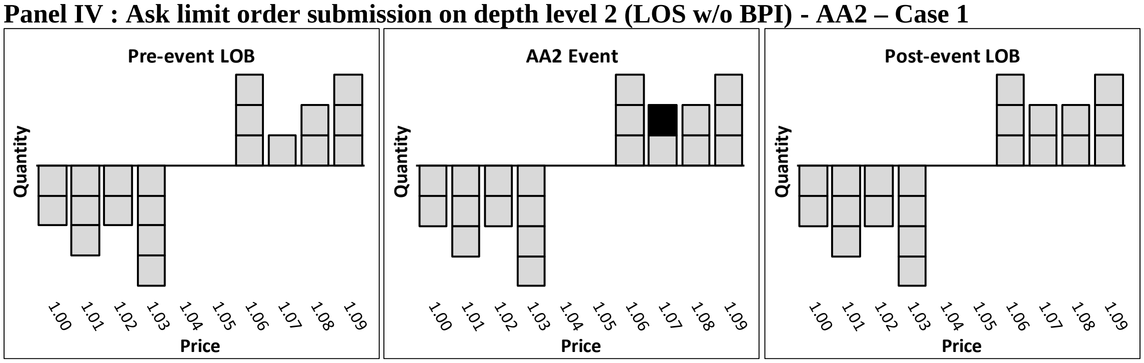
Panel II : Buy trade without best price impact (Trade w/o BPI) - Buy



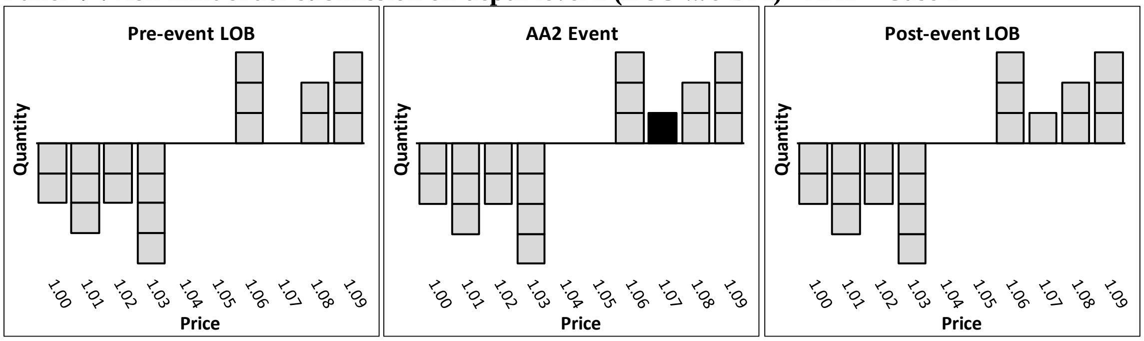
Panel III : Ask limit order submission with best price impact (LOS w/ BPI) - AA1*



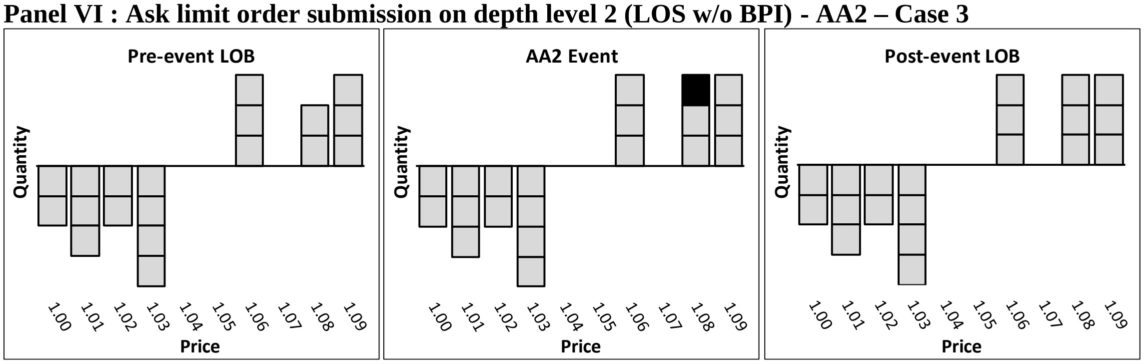
Panel IV : Ask limit order submission on depth level 2 (LOS w/o BPI) - AA2 – Case 1



Panel V : Ask limit order submission on depth level 2 (LOS w/o BPI) - AA2 – Case 2



Panel VI : Ask limit order submission on depth level 2 (LOS w/o BPI) - AA2 – Case 3



Panel VII : Bid limit order cancellation on depth level 1 (LOC) - BC1

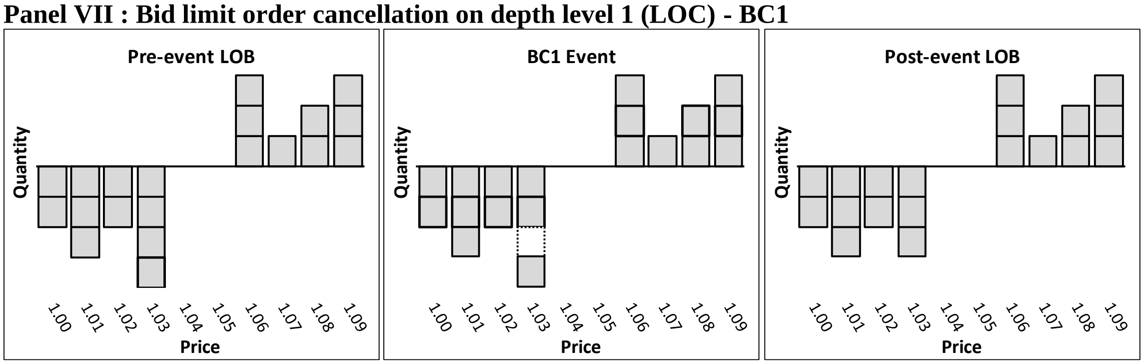


Figure 4.3 Events occurrences count daily averages

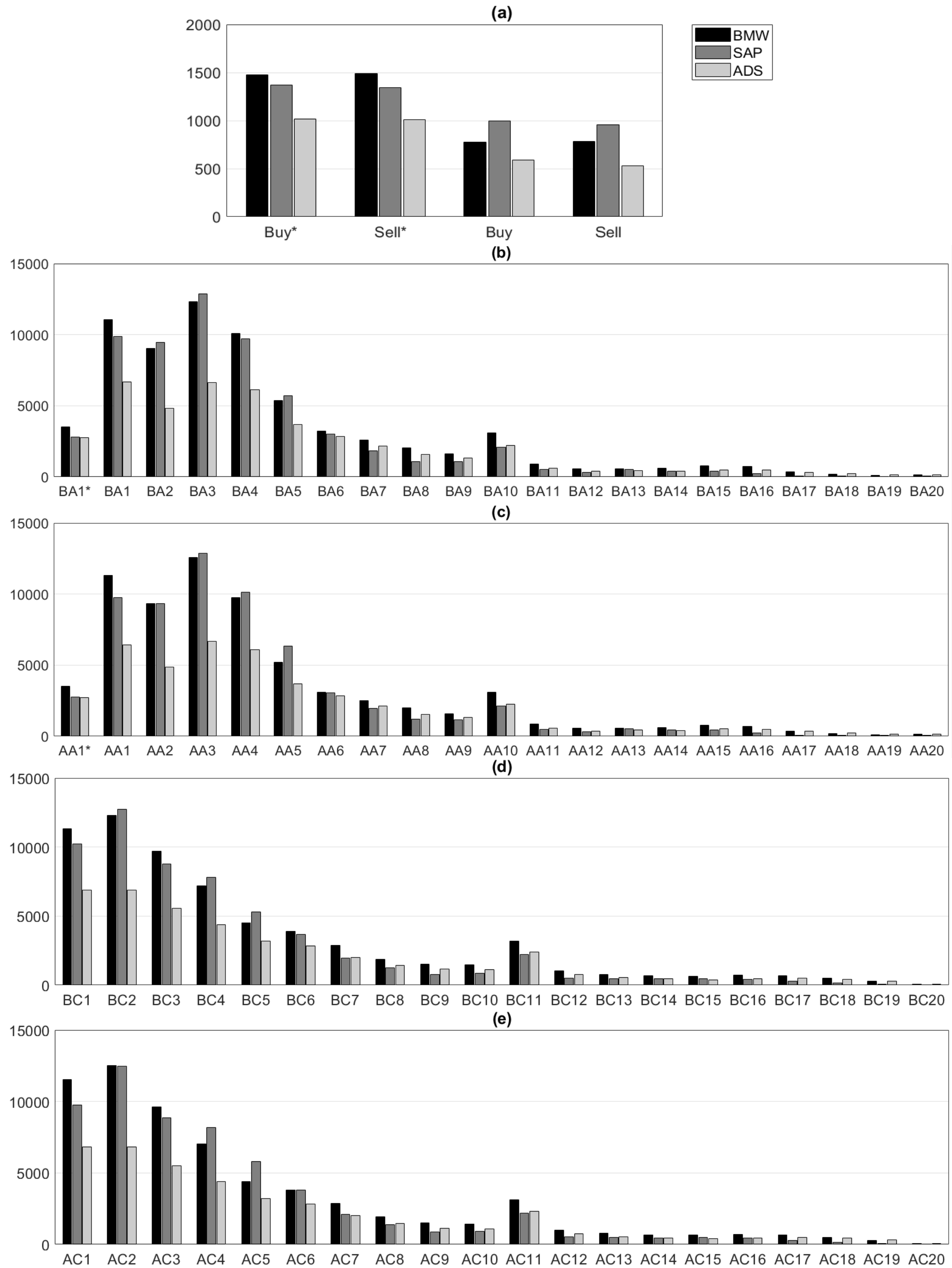
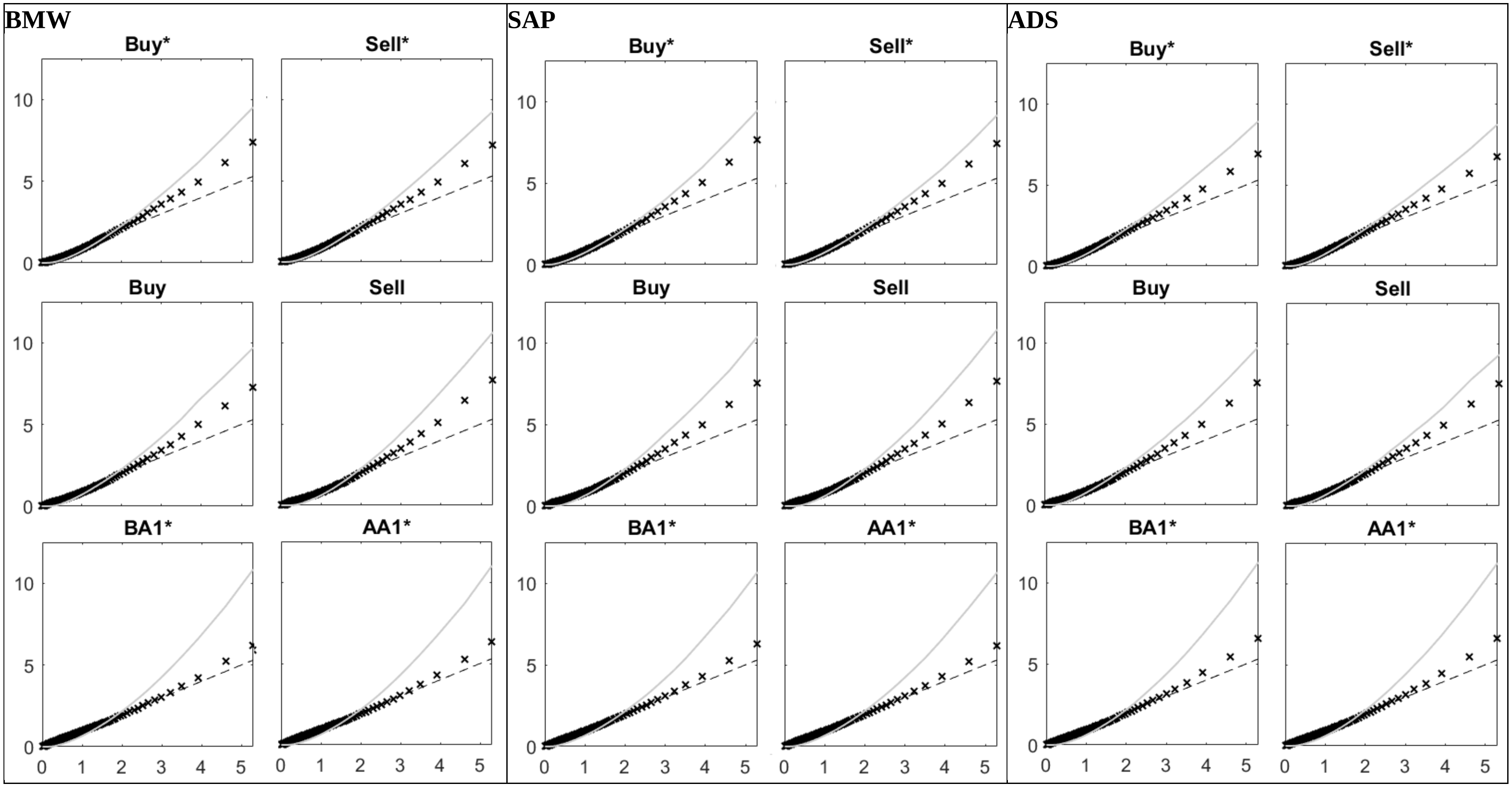


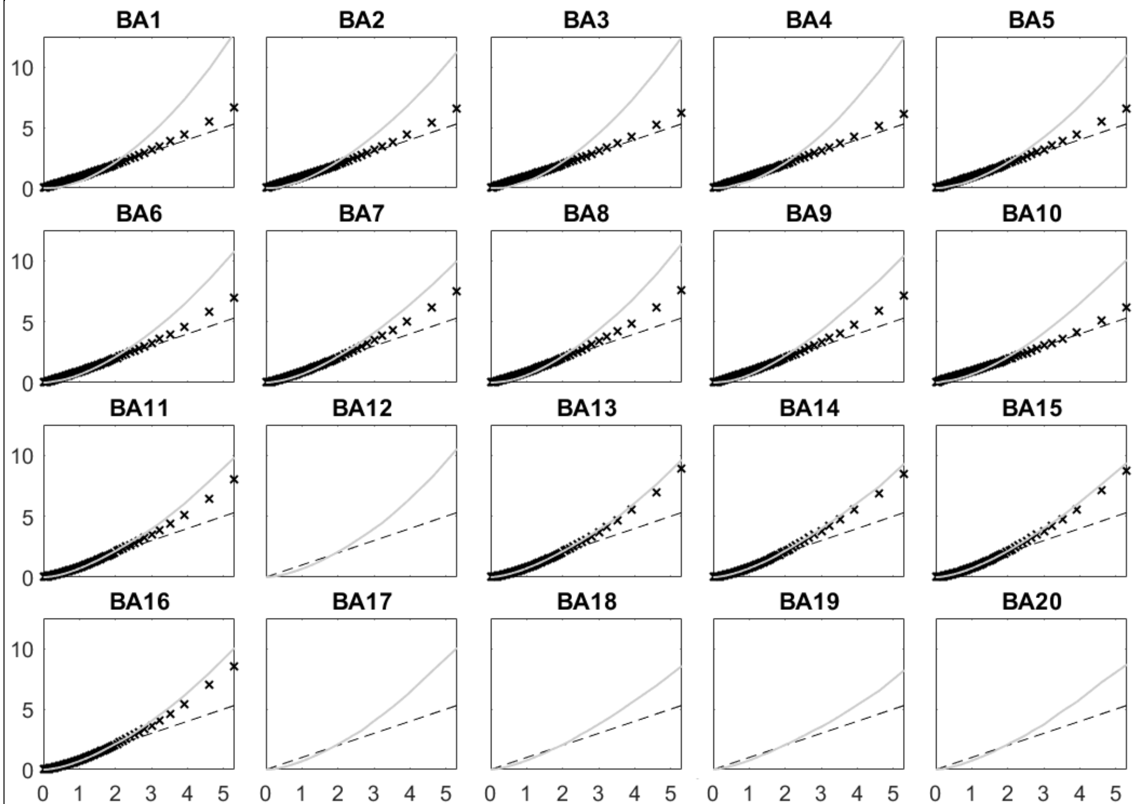
Figure 4.4 Descriptive models Q-Q Plots

Panel I: Trades w/ BPI, Trades w/o BPI and LOS w/ BPI

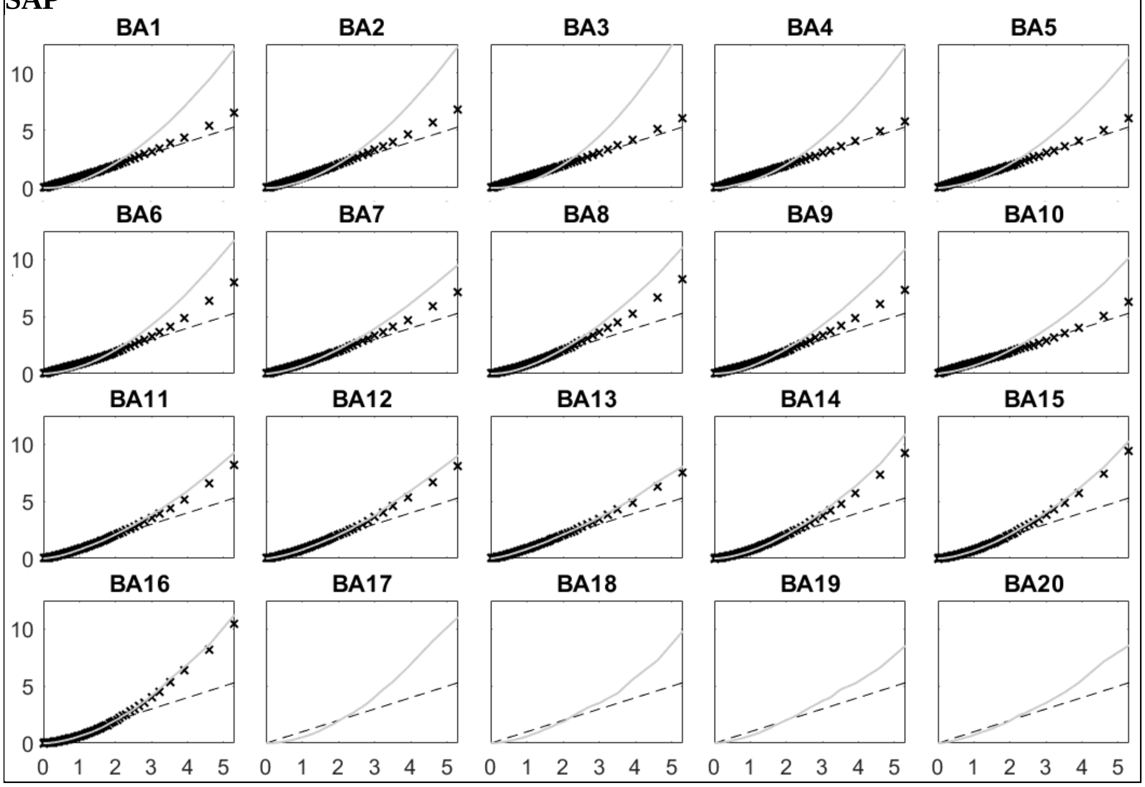


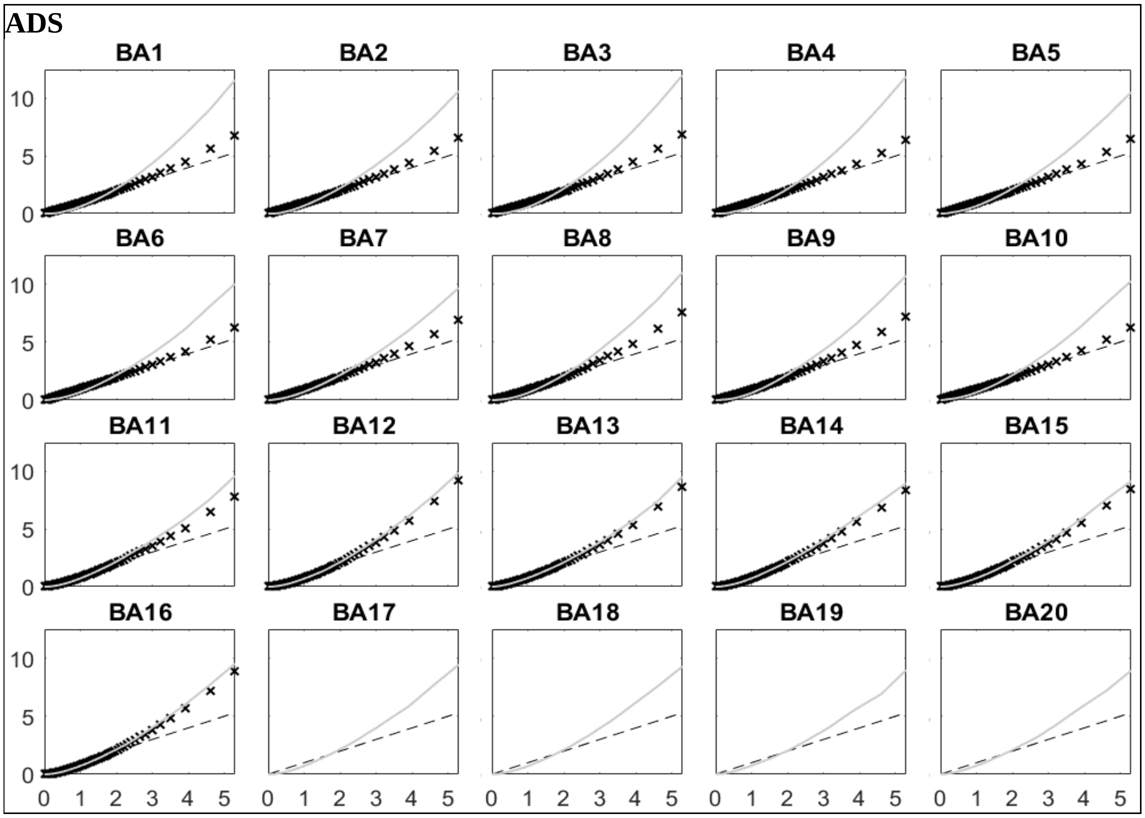
Panel II: Bid side LOS w/o BPI

BMW

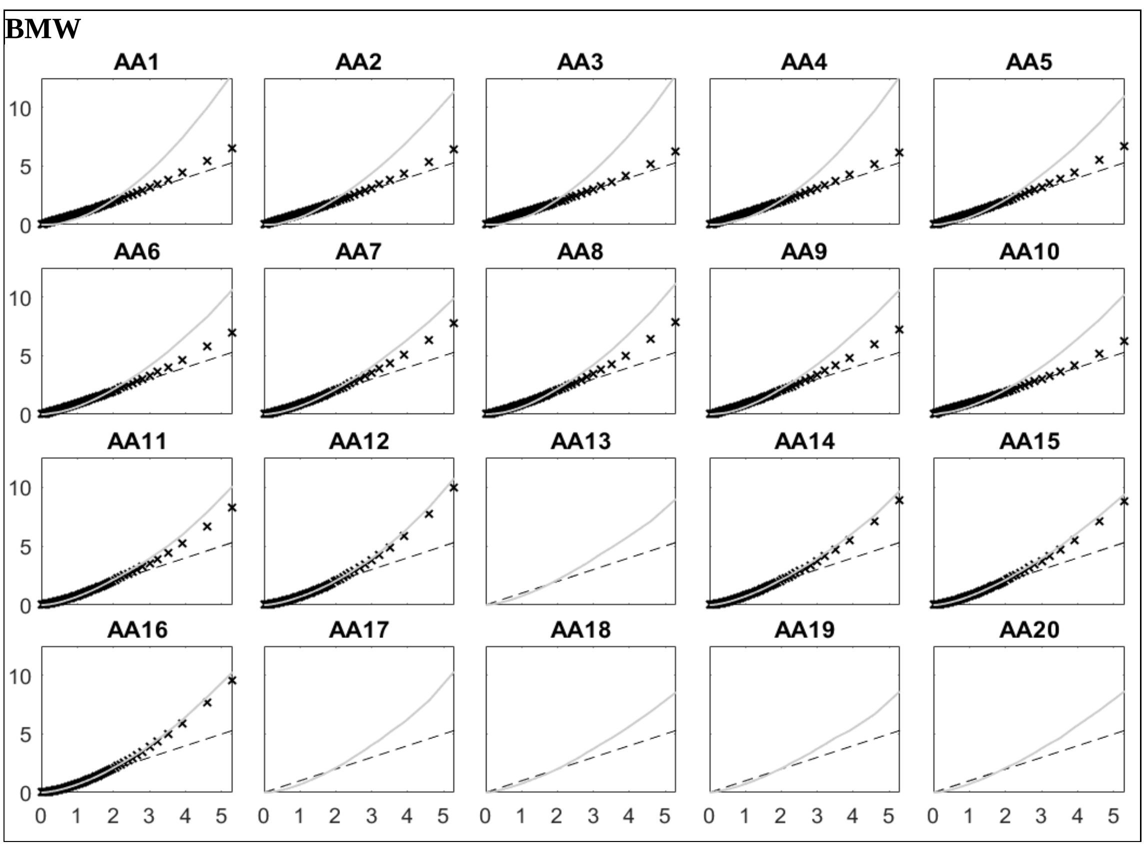


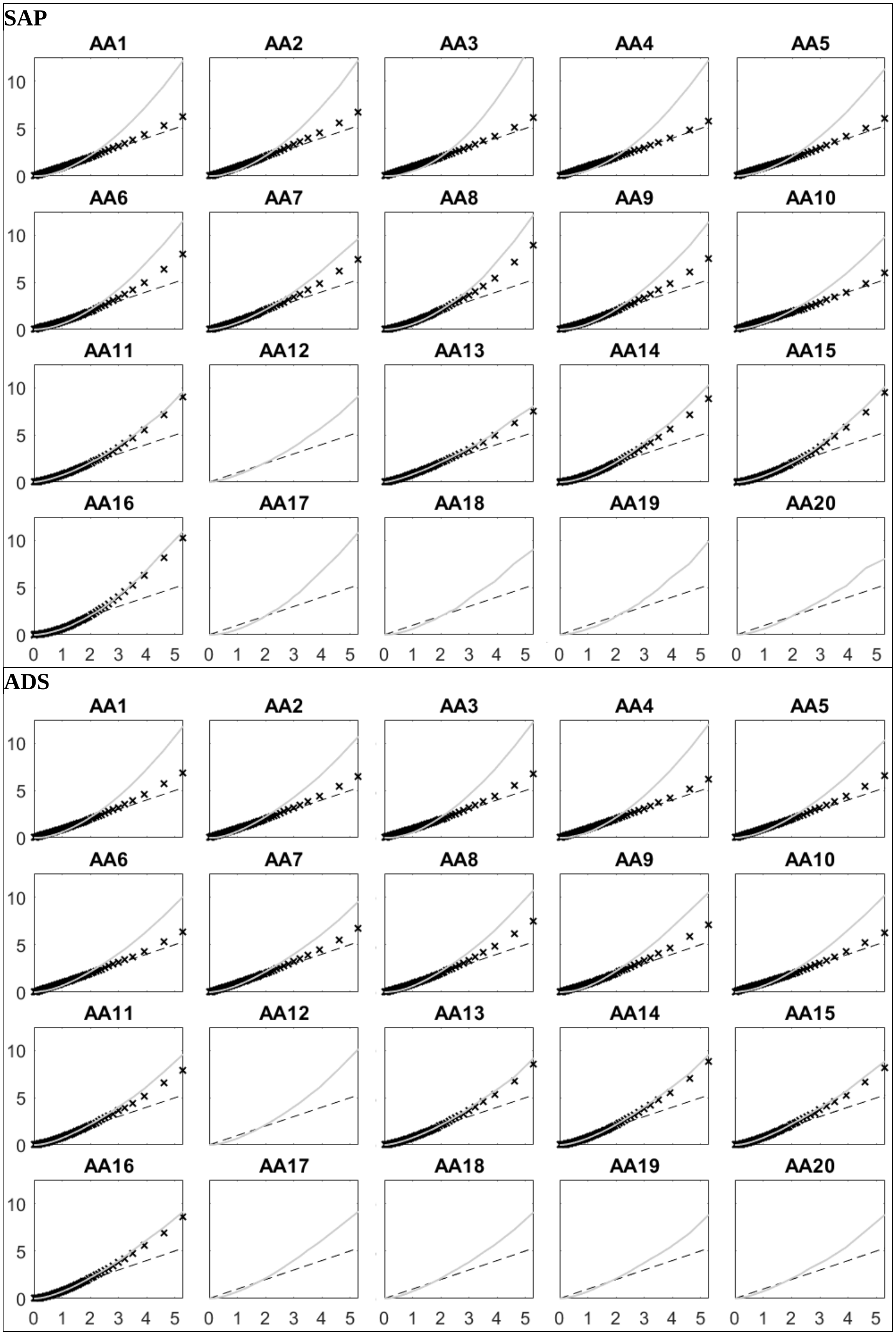
SAP



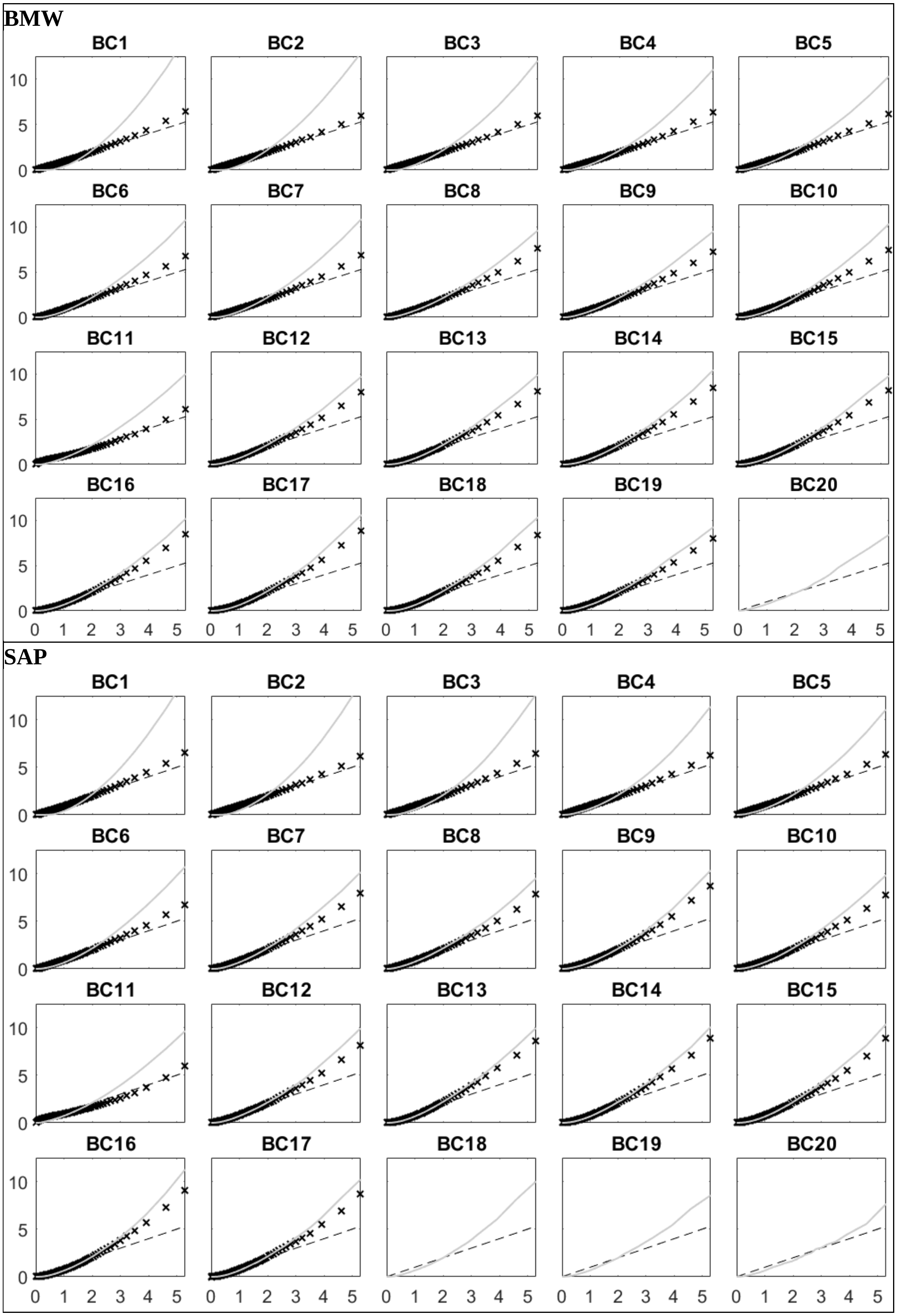


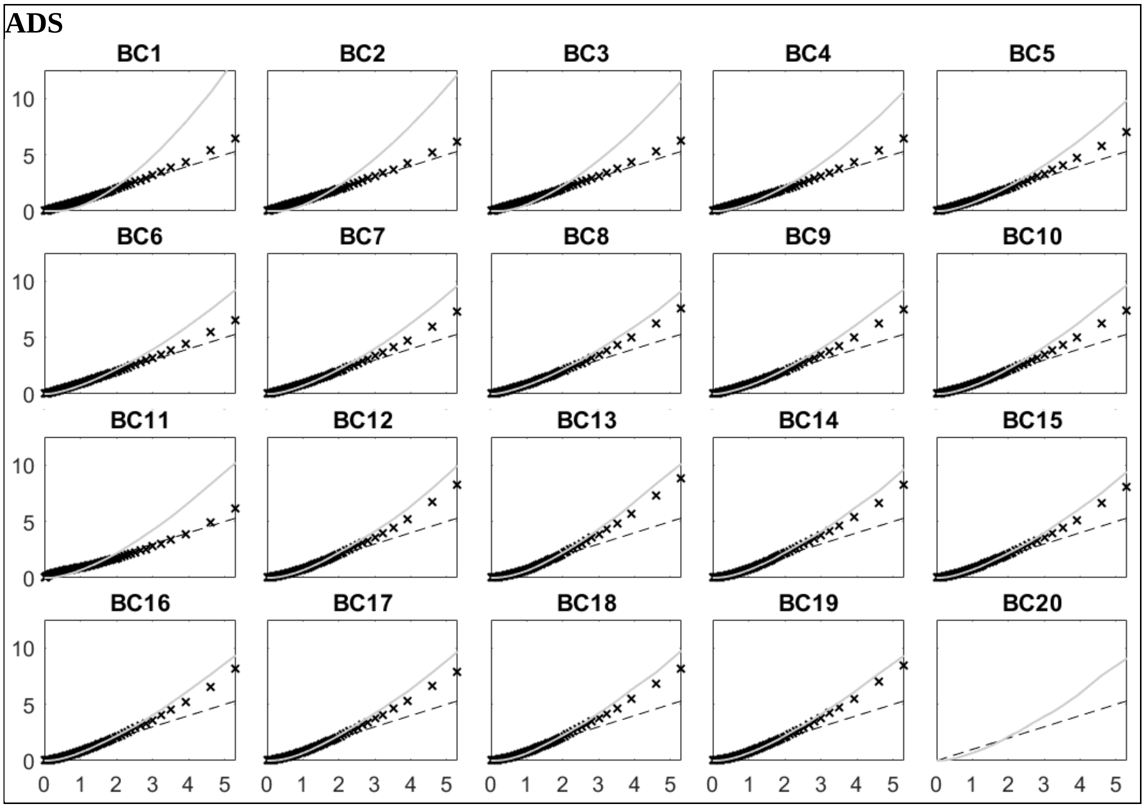
Panel III: Ask side LOS w/o BPI



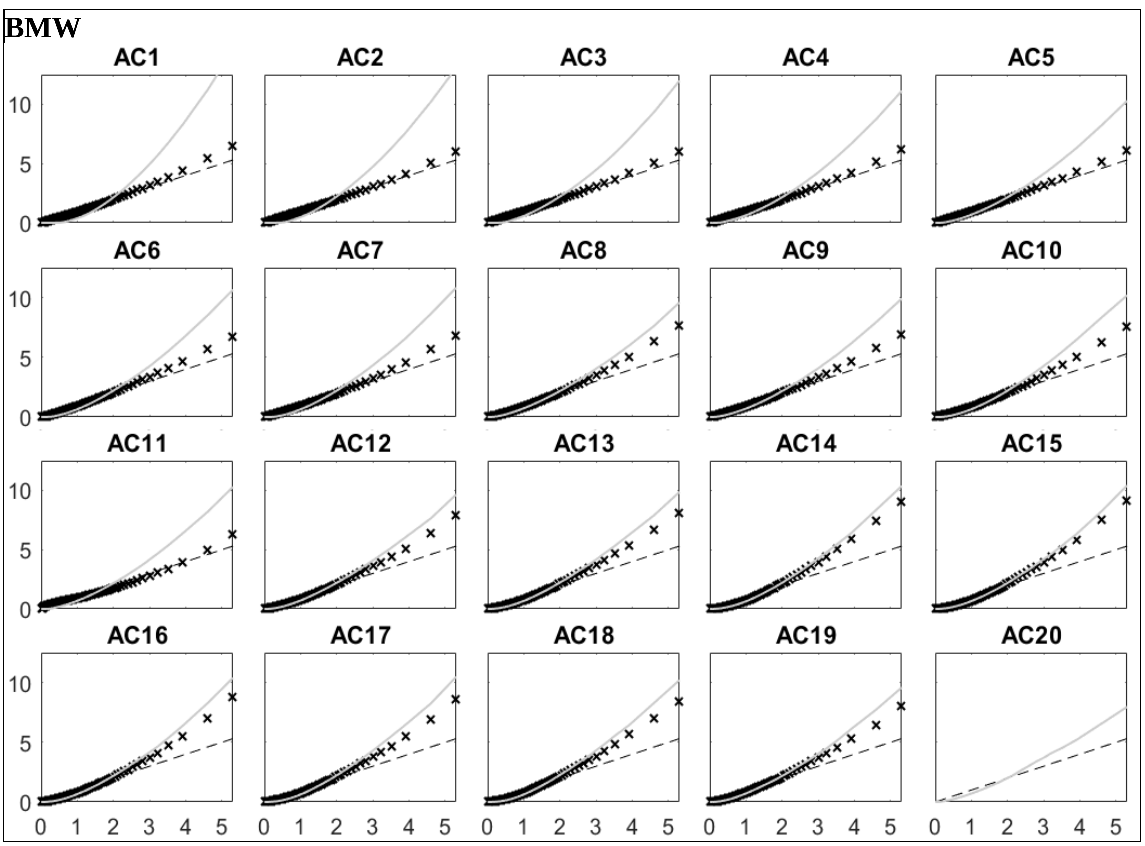

SAP
AA1
AA2
AA3
AA4
AA5
AA6
AA7
AA8
AA9
AA10
AA11
AA12
AA13
AA14
AA15
AA16
AA17
AA18
AA19
AA20
ADS
AA1
AA2
AA3
AA4
AA5
AA6
AA7
AA8
AA9
AA10
AA11
AA12
AA13
AA14
AA15
AA16
AA17
AA18
AA19
AA20

Panel IV: Bid side LOC





Panel V: Ask side LOC



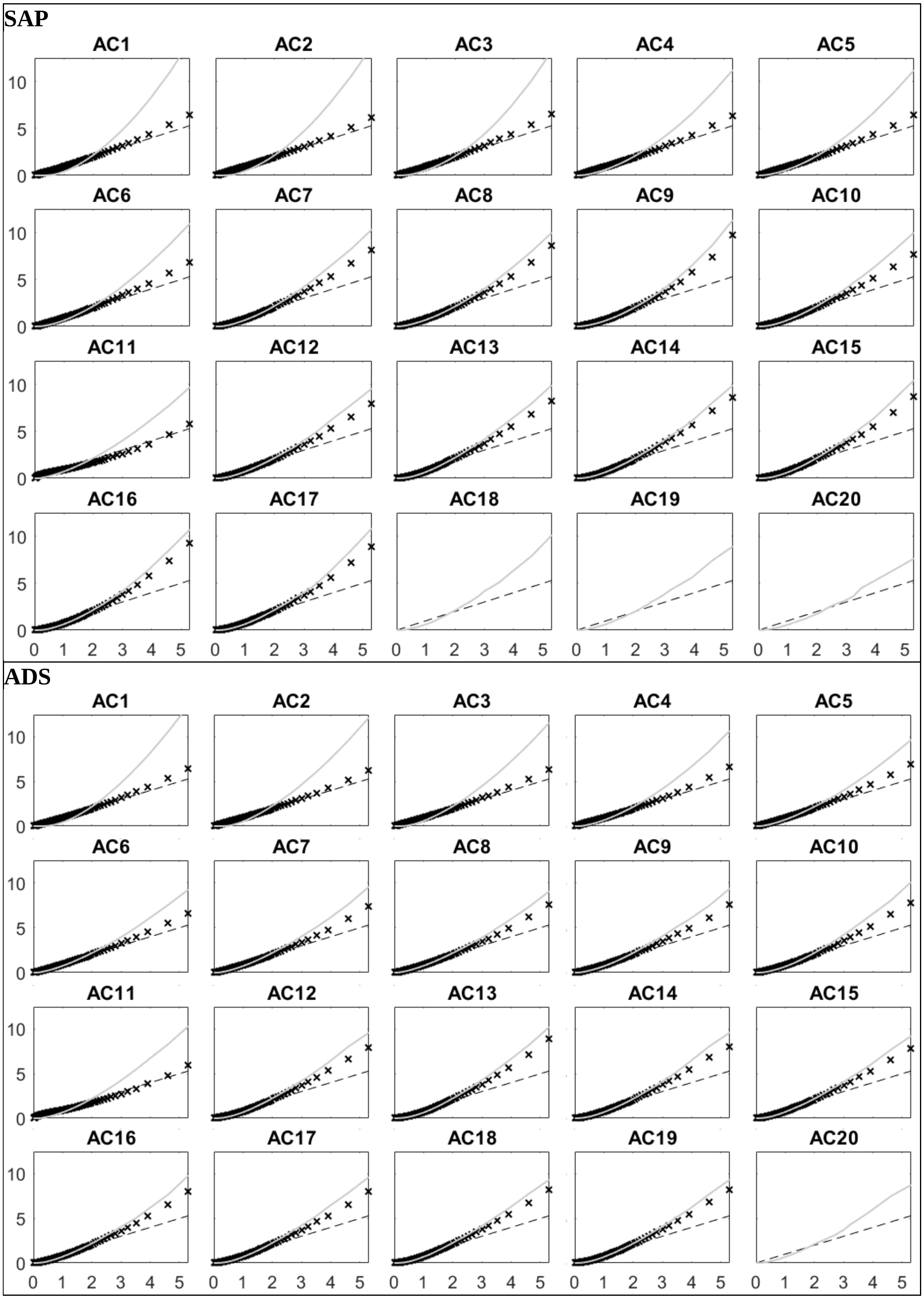

SAP
AC1
AC2
AC3
AC4
AC5
AC6
AC7
AC8
AC9
AC10
AC11
AC12
AC13
AC14
AC15
AC16
AC17
AC18
AC19
AC20
ADS
AC1
AC2
AC3
AC4
AC5
AC6
AC7
AC8
AC9
AC10
AC11
AC12
AC13
AC14
AC15
AC16
AC17
AC18
AC19
AC20

Figure 4.5 Events arrival process adjusted baseline daily average

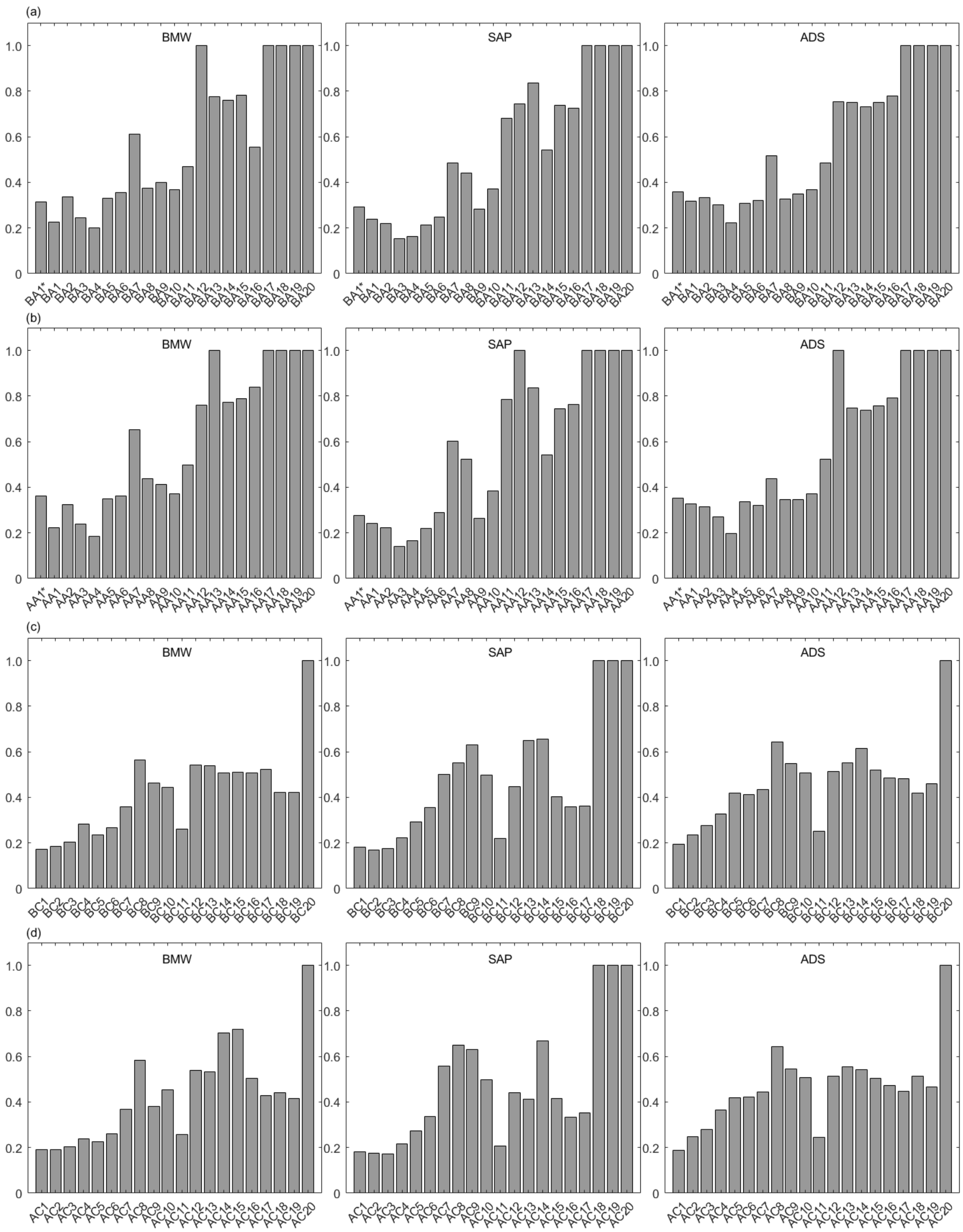
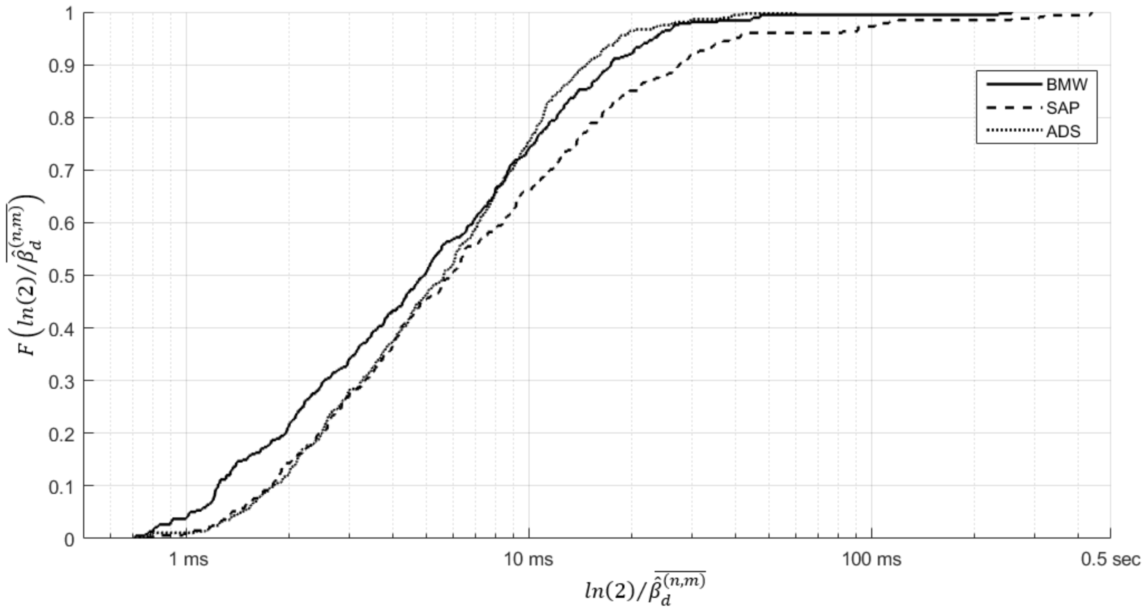
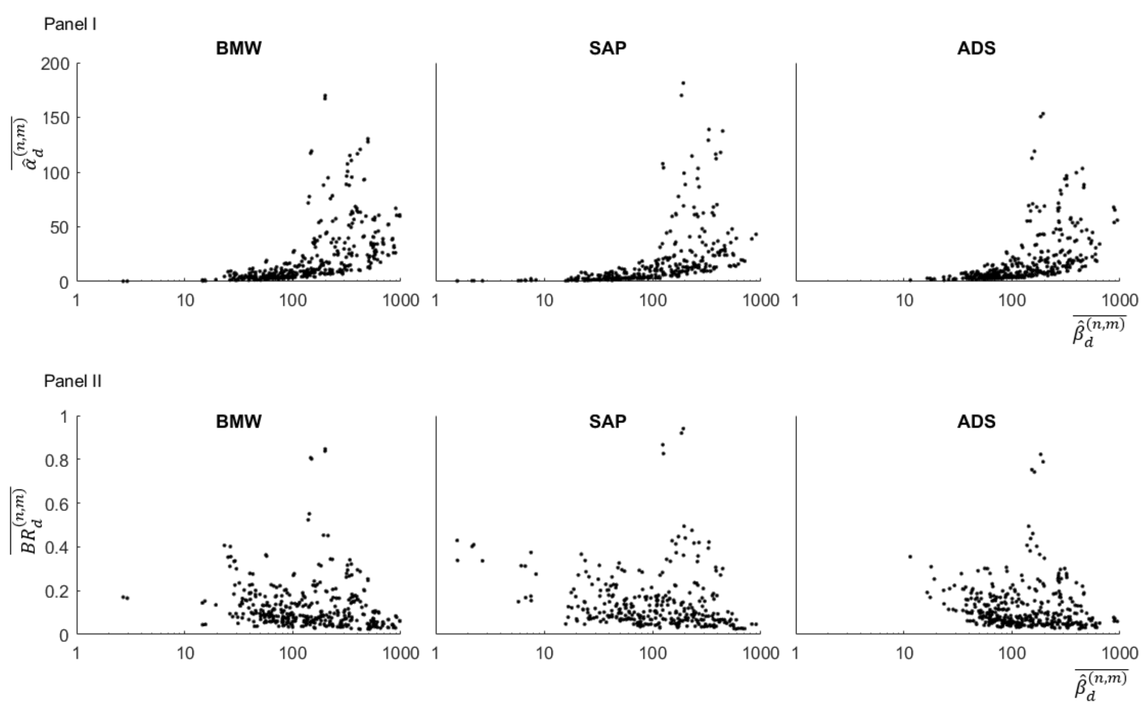


Figure 4.6 Hawkes effects half-life daily average CDF



This figure presents the cumulative distribution functions of the descriptive models Hawkes effects half-life daily averages.

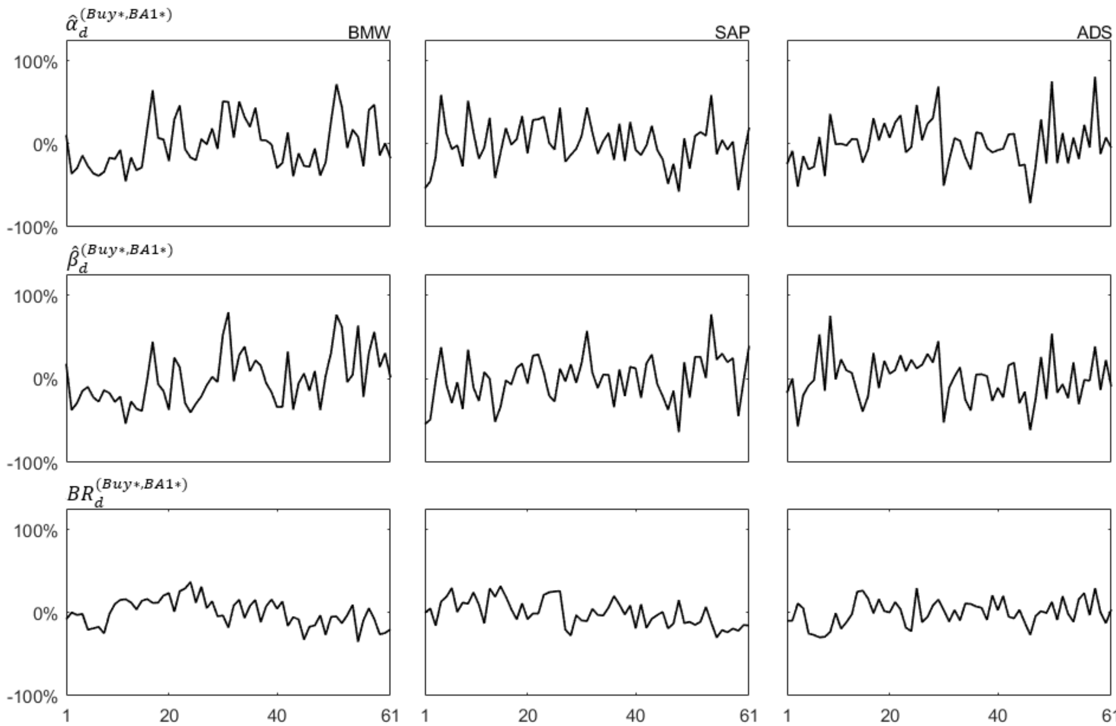
Figure 4.7 Hawkes effects Alpha parameters and Branching ratio with respect to Beta



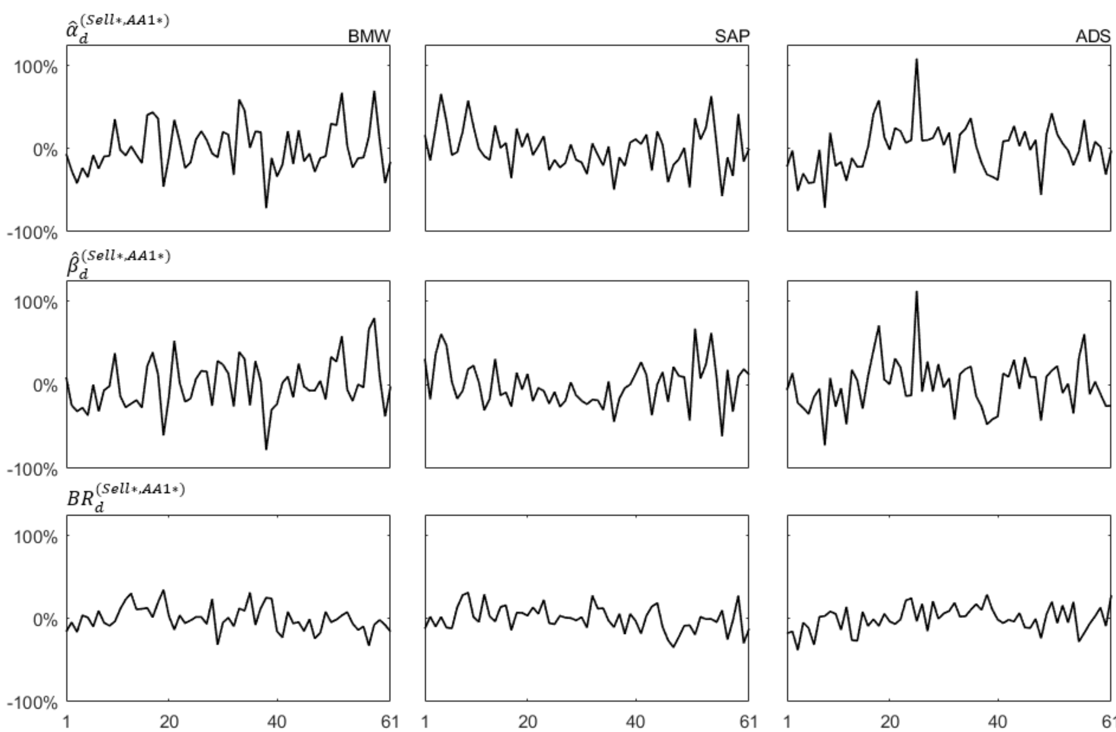
Panel I : $\hat{\alpha}_d^{(n,m)}$ daily average with respect to $\hat{\beta}_d^{(n,m)}$ daily average. Panel II : $BR_d^{(n,m)}$ daily average with respect to $\hat{\beta}_d^{(n,m)}$ daily average.

Figure 4.8 Daily estimated parameters dispersion examples

Panel I: (Buy*, BA1*)



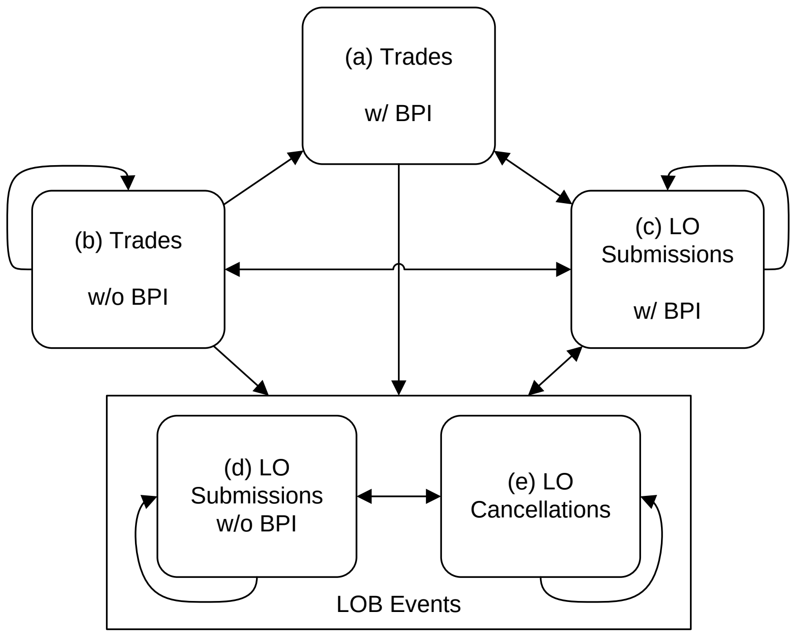
Panel II: (Sell*, AA1*)



This figure presents the daily dispersion with regard to the average for the complete models estimated parameters $\hat{\alpha}_d^{(buy*,BA1*)}$, $\hat{\alpha}_d^{(sell*,AA1*)}$, $\hat{\beta}_d^{(buy*,BA1*)}$ and $\hat{\beta}_d^{(sell*,AA1*)}$ as well as for the branching ratios obtained using these parameters.

Figure 4.9 General events dynamics

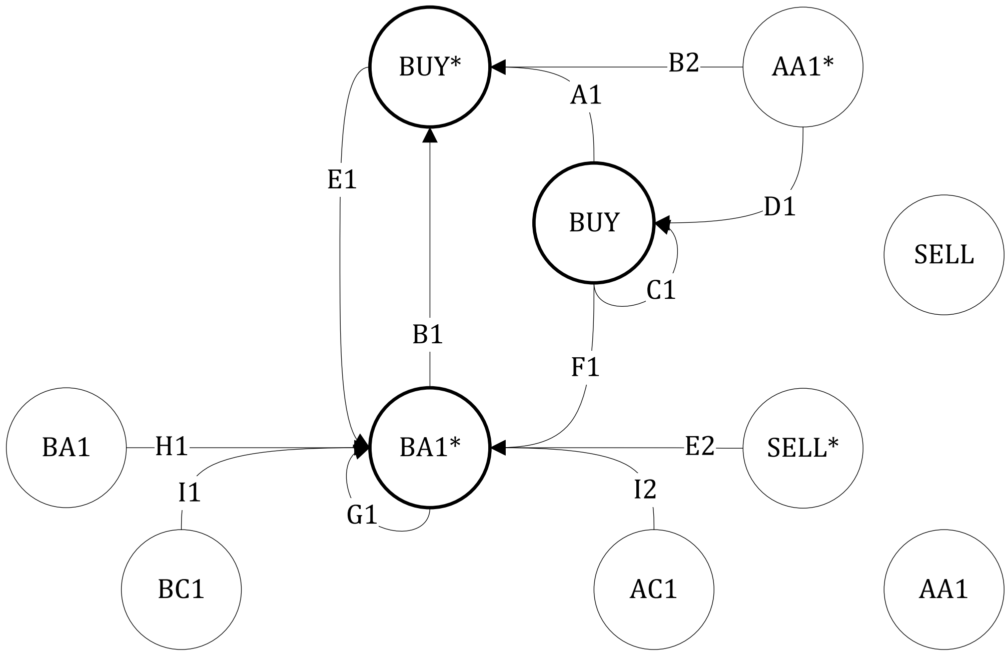
Panel I: High-level dynamics



Panel II: Events relationships classes

	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Bid LOS w/o BPI	Ask LOS w/o BPI	Bid LOC	Ask LOC
Buy*					E1	E2	J1	J2	K1	K2
Sell*					E2	E1	J2	J1	K2	K1
Buy	A1		C1		F1		L1			M1
Sell		A1		C1		F1		L1	M1	
BA1*	B1	B2		D1	G1		N1		O1	O2
AA1*	B2	B1	D1			G1		N1	O2	O1
Bid LOS w/o BPI					H1		P1		Q1	Q2
Ask LOS w/o BPI						H1		P1	Q2	Q1
Bid LOC					I1	I2	R1	R2	S1	
Ask LOC					I2	I1	R2	R1		S1

Panel III: Buyer side Trades w/ BPI, Trades w/o BPI and LOS w/ BPI



Panel IV: Buyer side LOS w/o BPI and LOC

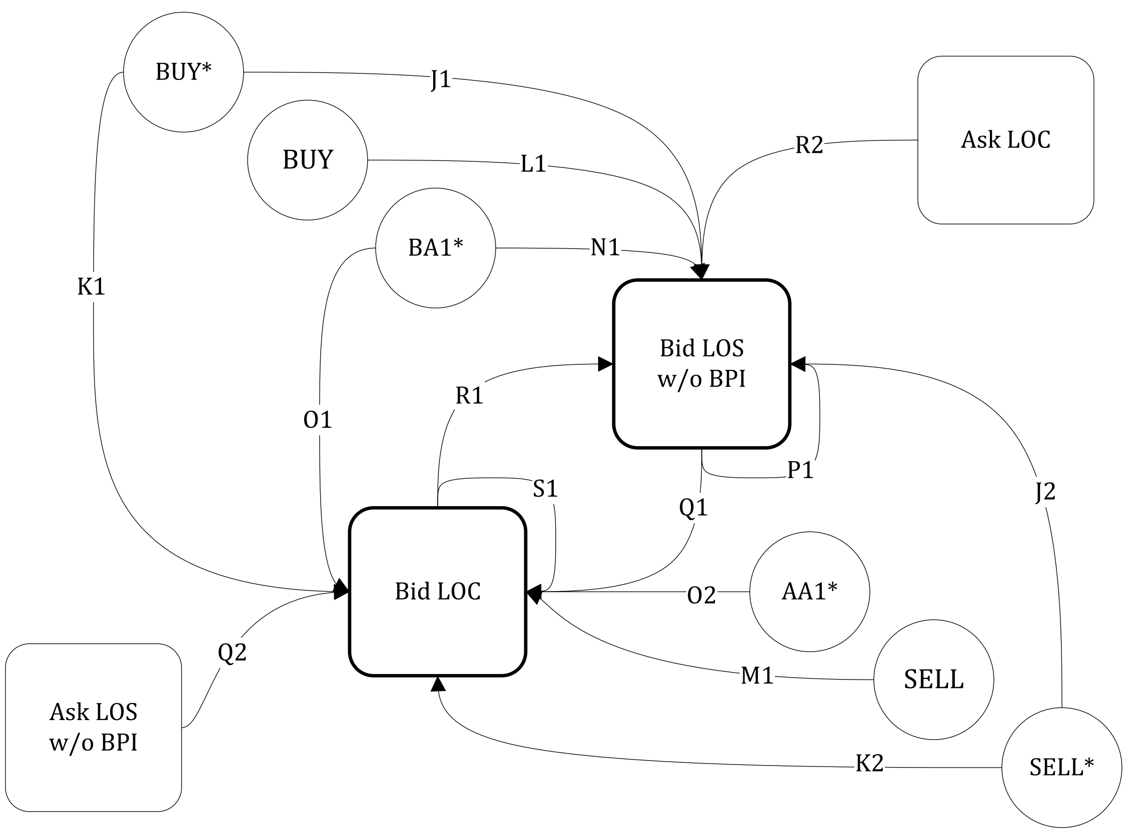
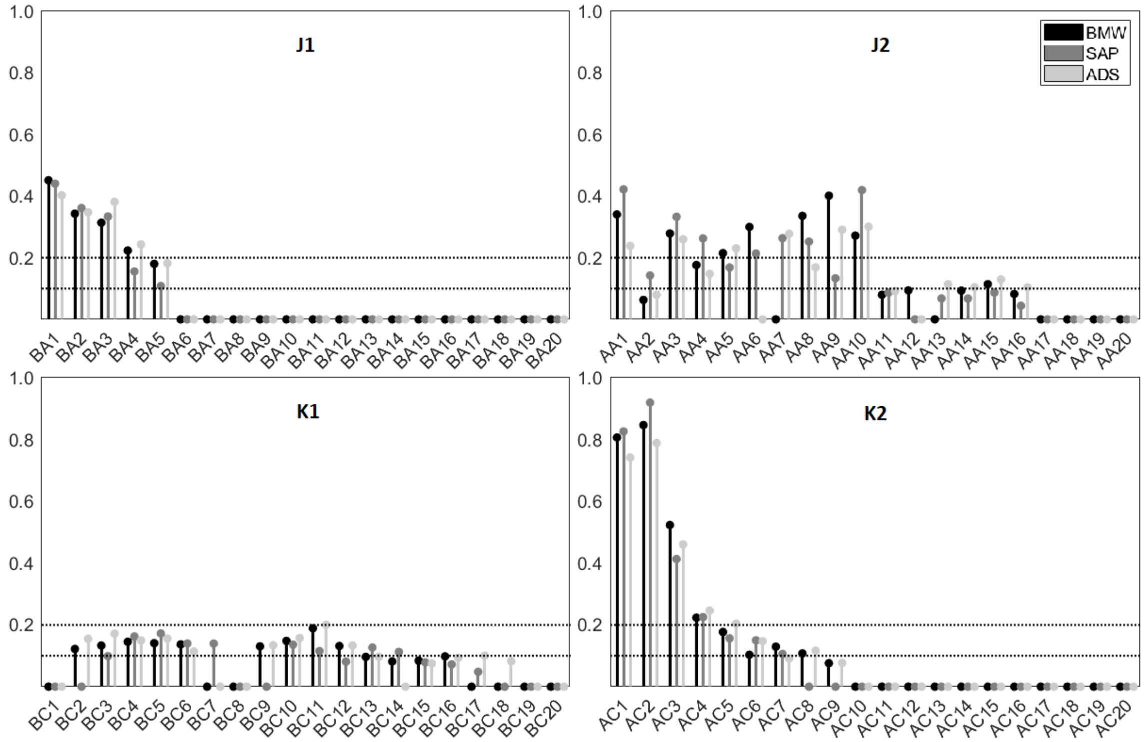
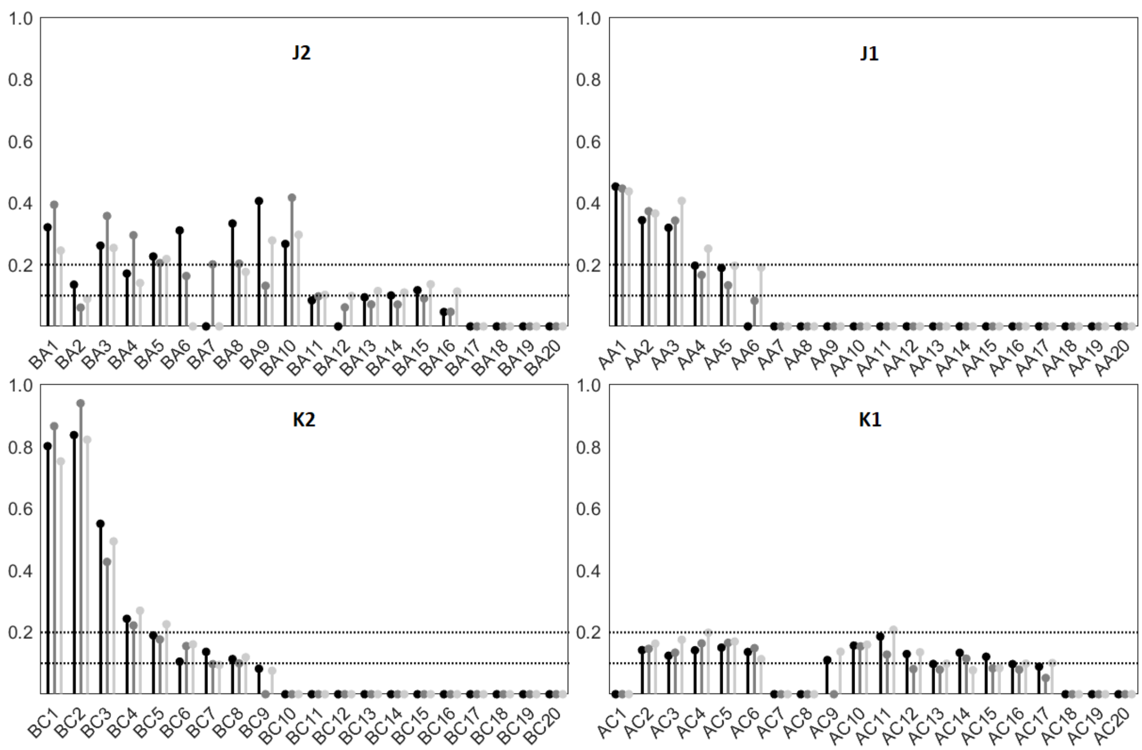


Figure 4.10 Trades and LOS w/ BPI effects branching ratio daily average

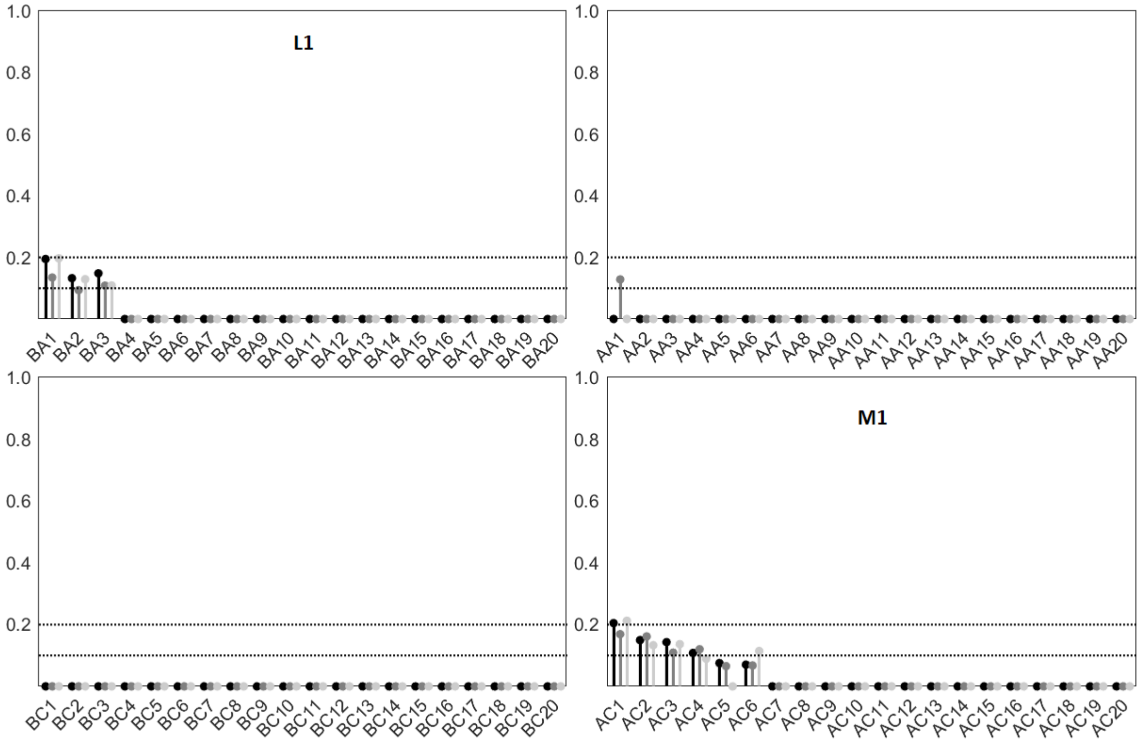
Panel I: Buy*



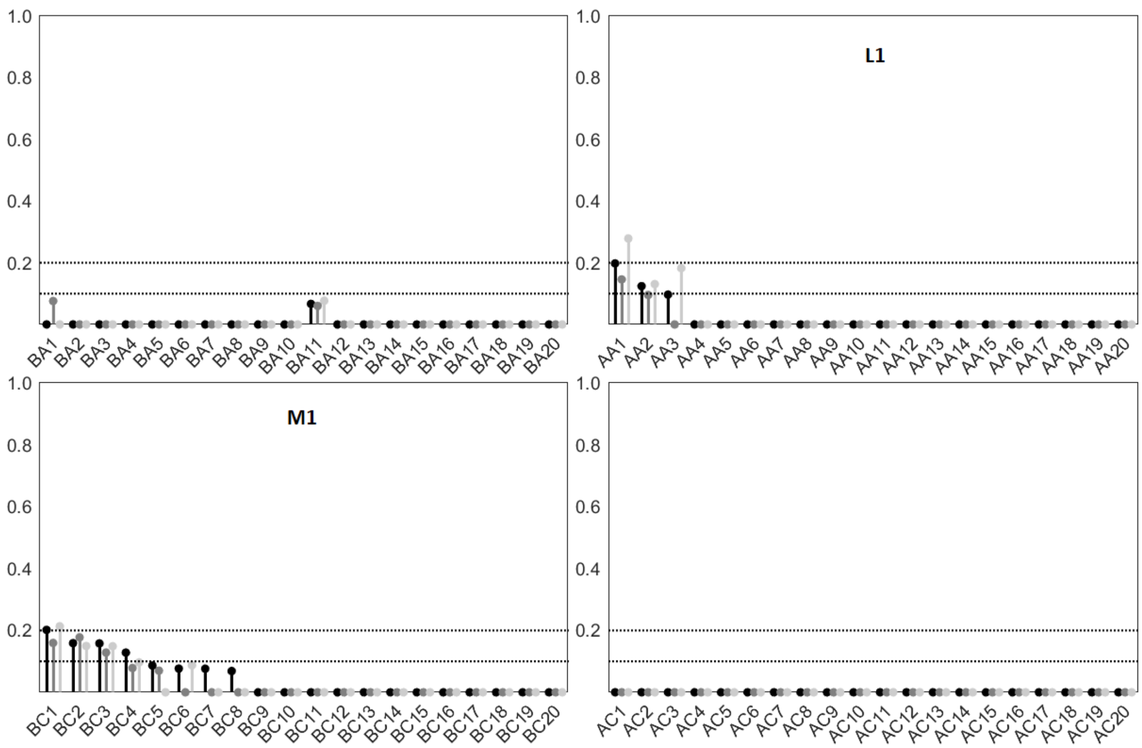
Panel II: Sell*



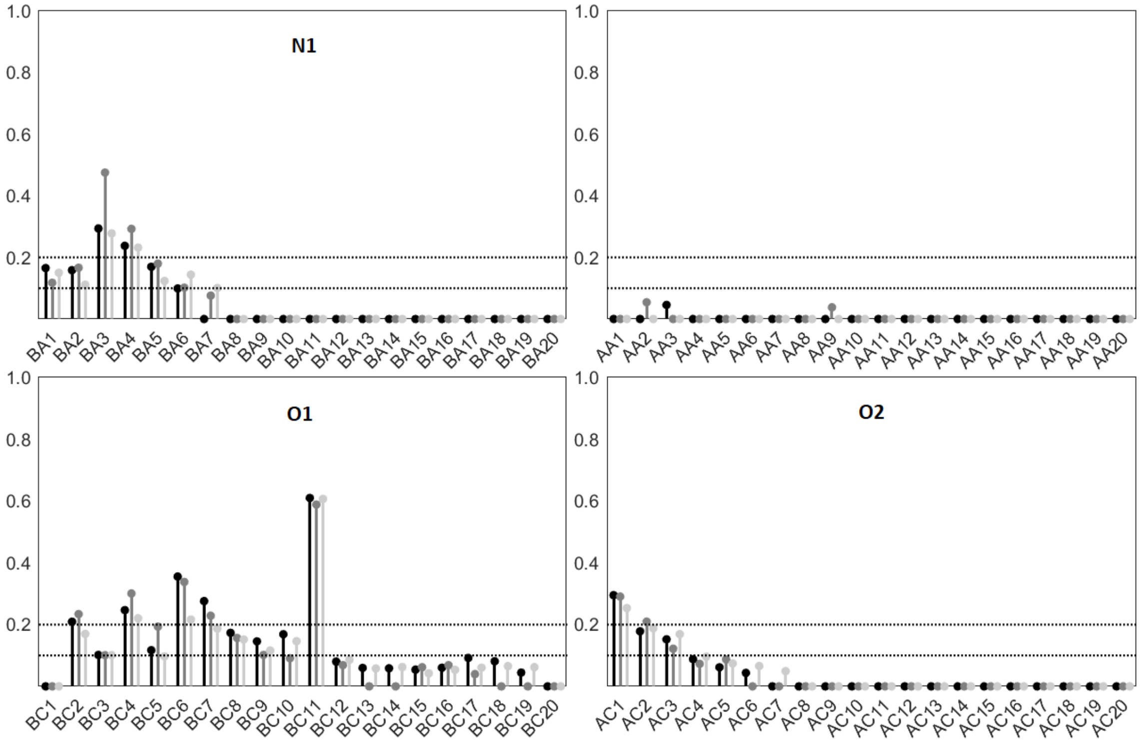
Panel III: Buy



Panel IV: Sell



Panel V: BA1*



Panel VI: AA1*

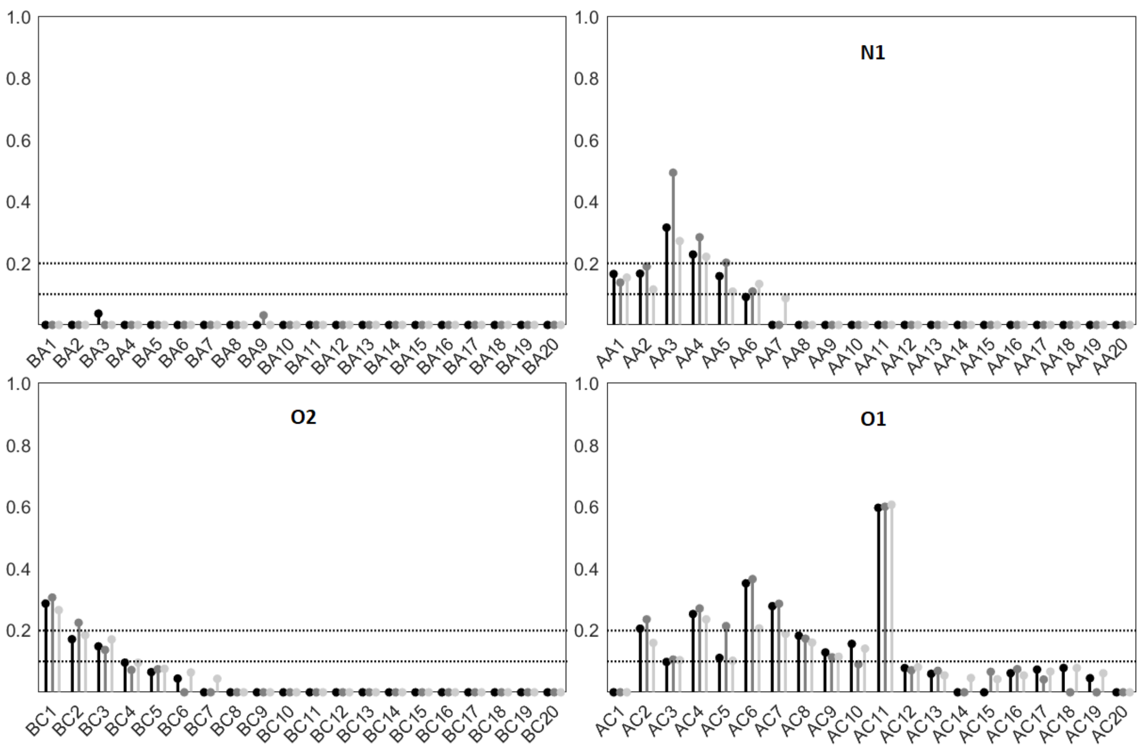


Figure 4.11 Category aggregated Hawkes effects contributions to LOB events arrival daily average

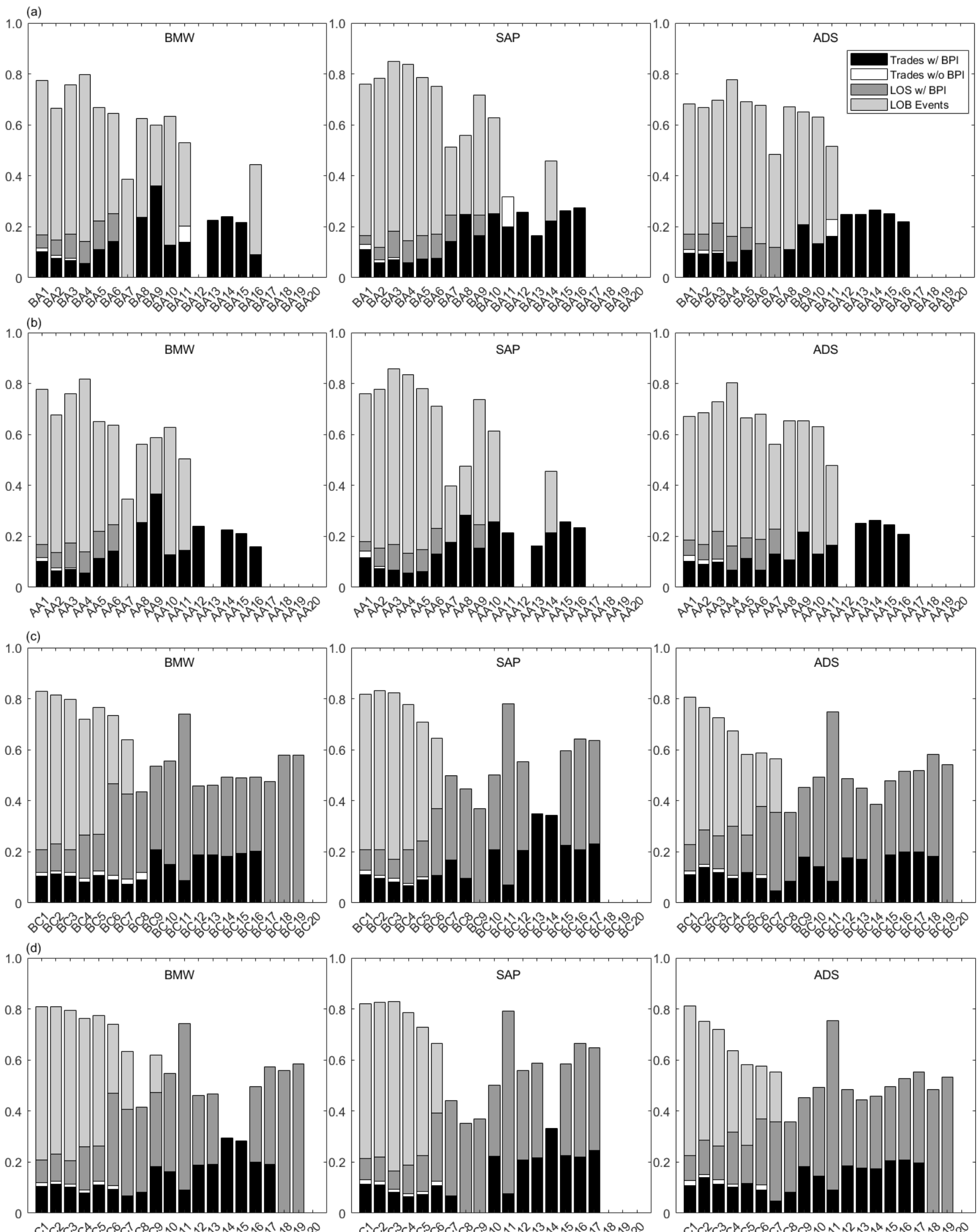


Table 4.1 Potential explanatory events sets definition

	Trades w/ BPI		Trades w/o BPI		LOS w/ BPI		Bid LOS				Ask LOS				Bid LOC				Ask LOC			
	Buy*	Sell*	Buy	Sell	BA1*	AA1*	BA1 … BA3	BA4 … BA6	BA7 … BA10	BA11 … BA20	AA1 … AA3	AA4 … AA6	AA7 … AA10	AA11 … AA20	BC1 … BC3	BC4 … BC6	BC7 … BC10	BC11 … BC20	AC1 … AC3	AC4 … AC6	AC7 … AC10	AC11 … AC11
Buy*	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Sell*	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Buy	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Sell	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
BA1*	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
AA1*	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
BA1 … BA3	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
BA4 … BA6	√	√	√	√	√	√		√	√	√		√	√	√		√	√	√		√	√	√
BA7 … BA10	√	√	√	√	√	√			√	√			√	√			√	√			√	√
BA11 … BA20	√	√	√	√	√	√				√				√				√				√
AA1 … AA3	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
AA4 … AA6	√	√	√	√	√	√		√	√	√		√	√	√		√	√	√		√	√	√
AA7 … AA10	√	√	√	√	√	√			√	√			√	√			√	√			√	√
AA11 … AA20	√	√	√	√	√	√				√				√				√				√
BC1 … BC3	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
BC4 … BC6	√	√	√	√	√	√		√	√	√		√	√	√		√	√	√		√	√	√
BC7 … BC10	√	√	√	√	√	√			√	√			√	√			√	√			√	√
BC11 … BC20	√	√	√	√	√	√				√				√				√				√
AC1 … AC3	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
AC4 … AC6	√	√	√	√	√	√		√	√	√		√	√	√		√	√	√		√	√	√
AC7 … AC10	√	√	√	√	√	√			√	√			√	√			√	√			√	√
AC11 … AC11	√	√	√	√	√	√				√				√				√				√

This table presents the potential events relationships tested through the estimation of our complete models. For each case presenting a check sign, the relationships between predecessor events (row) and the successor events (column) are tested.

Table 4.2 Descriptive Model Selection

Panel I: Trades, BA1* and AA1*

	BMW						SAP						ADS					
	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*
Buy*	6	0	11	11	61	55	2	1	2	19	61	54	4	1	7	17	61	50
Sell*	1	2	8	7	52	61	1	1	14	4	57	61	1	1	10	8	53	58
Buy	59	0	56	2	51	0	61	0	60	3	59	0	56	0	52	1	48	1
Sell	0	59	2	57	0	57	0	61	1	59	0	58	0	52	1	47	1	51
BA1*	45	61	4	44	49	5	25	61	5	45	21	3	50	60	8	26	55	5
AA1*	61	36	49	5	2	50	59	17	54	13	4	17	57	51	30	11	4	56
BA1	1	0	0	0	59	0	1	0	0	0	61	0	3	0	0	0	58	0
BA2	0	0	0	0	7	0	0	0	0	0	1	0	1	0	0	0	27	0
BA3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
BA4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA9	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
BA10	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0
BA11	0	0	0	0	0	1	0	0	0	0	0	0	0	5	0	0	0	0
BA12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA13	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
BA14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
BA15	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
BA16	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
BA17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA19	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
BA20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
AA1	0	2	0	0	0	61	0	3	0	0	0	60	0	8	0	0	0	58
AA2	0	1	0	0	0	8	0	0	0	0	0	6	0	0	0	0	0	20
AA3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	4
AA4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
AA5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
AA8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
AA9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA10	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
AA11	3	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
AA12	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
AA13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA14	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
AA15	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
AA16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
AA17	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
AA18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

This table presents the number of trading days for which a given events relationship meets our selection criteria. Each column represents an explained event and each row, an explanatory event. The tone scale relates to the following categories: white: 0 to 6 days; pale grey: 7 to 30 days; dark grey: 31 to 54 days; black: 55 to 61 days.

Panel I: Trades, BA1* and AA1* (cont.)

	BMW						SAP						ADS					
	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*
BC1	0	0	0	0	34	61	0	0	2	0	36	56	1	0	1	0	54	61
BC2	0	0	0	0	1	6	0	2	0	0	0	2	0	1	0	0	5	9
BC3	0	10	0	0	0	4	0	6	0	0	0	1	1	13	0	0	0	7
BC4	0	15	0	0	0	1	0	0	0	1	0	0	1	4	0	0	0	1
BC5	0	6	0	1	0	1	0	0	0	0	0	0	0	5	0	0	0	5
BC6	0	1	0	1	0	0	0	3	0	1	0	0	1	5	0	2	0	1
BC7	0	1	0	0	0	0	0	2	0	0	0	0	0	8	0	0	0	0
BC8	0	0	0	0	0	0	0	0	0	1	0	0	0	4	0	1	0	1
BC9	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
BC10	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	2	0	0
BC11	2	2	1	3	0	0	1	5	1	8	0	1	2	2	4	2	1	0
BC12	2	0	0	0	0	0	1	0	0	1	0	0	0	1	1	0	0	0
BC13	2	0	0	0	2	0	0	1	0	0	0	0	1	0	0	0	0	0
BC14	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
BC15	2	1	0	0	3	0	0	0	0	0	1	0	0	0	0	0	1	0
BC16	1	0	1	0	4	0	1	0	1	0	0	1	1	0	0	1	1	1
BC17	0	1	0	1	2	1	1	0	0	0	0	0	2	0	0	0	0	1
BC18	1	0	0	0	2	0	0	0	1	0	1	0	0	0	0	0	1	0
BC19	0	1	0	0	2	0	0	0	0	0	0	0	2	0	0	1	0	0
BC20	0	0	0	0	1	0	1	2	0	0	0	0	0	0	0	0	3	1
AC1	0	0	0	0	60	28	0	0	0	0	58	34	0	0	0	0	61	56
AC2	0	0	0	0	6	0	2	0	0	0	5	0	0	0	0	0	12	1
AC3	9	0	1	0	11	0	1	0	0	0	0	0	8	0	0	0	8	1
AC4	10	0	0	0	1	0	0	0	0	0	0	0	7	0	0	0	4	0
AC5	8	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	4	1
AC6	1	0	0	0	0	0	1	0	0	0	0	0	12	0	0	0	0	0
AC7	3	0	1	0	0	0	4	0	1	0	1	0	6	0	0	0	2	1
AC8	1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0
AC9	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
AC10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
AC11	1	2	3	1	1	1	2	2	7	1	2	0	4	2	6	1	0	0
AC12	0	1	0	0	1	0	0	0	0	1	0	2	0	1	0	0	0	2
AC13	0	2	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
AC14	0	3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1
AC15	0	1	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	1
AC16	0	5	0	0	0	3	0	0	0	0	0	0	0	1	0	0	1	1
AC17	0	0	0	1	0	2	0	0	0	0	0	1	0	0	0	0	0	0
AC18	2	1	1	0	0	1	0	2	0	0	1	0	0	0	0	1	0	2
AC19	1	1	1	0	1	1	0	3	0	0	0	0	1	0	0	0	0	1
AC20	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0

Panel II: BA1 to BA10

	BMW										SAP										ADS									
BA	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Buy*	61	61	60	59	48	26	22	10	1	0	60	61	52	44	47	25	16	14	1	1	60	61	61	60	47	28	17	21	2	0
Sell*	61	39	61	54	44	34	30	42	60	61	61	33	55	40	39	46	38	45	54	60	61	31	55	53	31	23	27	44	61	61
Buy	55	59	58	27	7	2	1	2	0	0	55	48	32	1	3	0	0	0	0	0	57	55	44	17	8	5	4	3	0	0
Sell	22	15	12	5	4	8	3	3	3	6	37	17	8	3	2	3	1	1	0	12	25	4	7	4	4	2	0	2	0	6
BA1*	61	61	61	60	53	59	28	4	0	1	59	59	60	58	48	42	32	4	1	12	59	61	61	61	49	51	44	9	1	1
AA1*	21	23	47	14	3	23	22	25	13	17	13	23	23	8	9	11	11	5	43	19	7	19	8	12	8	12	23	21	24	20
BA1	61	61	61	57	14	0	1	0	0	0	61	60	59	13	8	1	0	0	0	2	56	61	61	47	23	2	1	1	0	0
BA2	47	59	59	21	5	0	1	0	0	0	39	19	50	13	4	0	0	0	0	0	29	54	55	11	2	3	2	0	0	0
BA3	26	59	60	35	6	1	1	0	0	0	4	51	40	41	17	0	1	0	0	0	18	43	55	43	23	16	10	1	0	0
BA4				25	51	3	3	0	0	0				7	38	0	0	0	0	0				7	32	9	4	0	0	0
BA5				49	17	47	6	1	2	3				3	0	29	0	0	0	0				29	8	43	4	1	0	0
BA6				13	24	2	42	10	1	14				3	3	1	32	3	0	2				3	38	1	45	5	0	5
BA7							0	52	7	9							0	43	2	5							1	53	4	6
BA8							9	1	58	7							7	0	39	5							6	1	55	9
BA9							10	32	1	59							1	12	0	53							3	28	1	58
BA10							9	4	0	0							4	2	0	0							3	2	0	0
AA1	1	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA2	0	1	7	2	1	1	1	0	0	0	0	0	1	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA3	0	0	49	11	0	0	0	0	0	0	0	1	22	6	3	0	0	0	0	0	0	0	0	1	0	3	1	0	0	0
AA4				27	0	0	0	0	0	0				7	0	0	0	0	0	0				18	0	0	0	0	0	0
AA5				8	8	2	1	0	0	0				0	3	0	0	0	0	0				1	4	0	1	0	0	0
AA6				4	2	2	0	0	0	0				4	4	1	0	0	0	0				1	0	0	0	0	0	0
AA7							0	0	0	0							2	0	0	0							0	0	0	0
AA8							1	2	0	0							0	0	0	0							0	0	0	0
AA9							2	0	0	0							1	0	0	1							1	0	0	0
AA10							0	0	0	0							0	0	1	9							1	0	0	0
BC1	61	55	61	29	24	18	12	23	2	59	61	48	60	31	8	35	13	19	0	56	61	61	53	32	26	10	29	45	19	58
BC2	34	42	60	37	1	2	24	14	24	0	39	45	56	49	13	2	13	3	52	0	48	55	60	42	5	5	26	41	58	0
BC3	21	59	60	59	23	4	11	10	0	59	12	38	61	56	39	1	2	1	0	2	21	54	30	61	20	5	7	19	0	60
BC4				46	51	17	5	5	0	0				61	36	36	0	0	0	0				28	50	14	4	4	0	0
BC5				50	57	19	4	2	8	0				29	59	23	8	0	1	0				60	57	33	3	2	10	0
BC6				55	11	57	0	4	0	1				43	14	61	11	0	0	0				59	26	59	9	1	0	8
BC7							42	6	0	0							43	5	0	0							43	15	4	1
BC8							0	43	14	3							8	37	1	2							7	39	16	7
BC9							2	4	19	3							3	1	8	0							5	3	18	5
BC10							4	0	1	12							1	0	2	4							3	2	8	18
AC1	61	61	61	59	28	0	0	0	0	0	61	58	58	41	3	0	0	0	0	0	61	61	61	61	53	4	0	0	0	0
AC2	54	57	59	22	1	0	0	0	0	0	44	51	52	7	0	0	0	0	0	0	33	56	56	39	3	1	0	0	0	0
AC3	57	60	61	46	3	0	1	0	0	0	29	33	54	3	1	0	0	0	0	0	37	53	59	36	9	0	0	0	0	0
AC4				26	3	0	1	0	0	0				3	2	2	0	0	0	0				31	6	6	2	0	0	0
AC5				8	0	0	0	0	0	0				2	8	1	4	0	0	0				10	6	5	5	1	0	0
AC6				19	0	0	0	1	0	0				5	3	1	6	0	0	0				20	0	0	3	1	0	0
AC7							1	7	2	0							1	5	0	0							0	1	4	0
AC8							0	1	1	2							1	0	0	0							1	0	0	3
AC9							1	1	0	0							0	1	0	2							1	3	0	1
AC10							6	5	1	0							7	7	3	3							1	1	1	1

Notes: Each column represents the depth level number on which the explained Bid order Added event occurs. Hatched cells correspond to relationship subject to the model initial restrictions described in section 2.5. The same applies to absent explanatory events.

Panel III: BA11 to BA20

	BMW										SAP										ADS									
BA	11	12	13	14	15	16	17	18	19	20	11	12	13	14	15	16	17	18	19	20	11	12	13	14	15	16	17	18	19	20
Buy*	0	0	2	6	0	1	2	0	0	12	2	2	25	12	0	1	0	0	0	0	0	0	0	0	1	3	6	3	2	10
Sell*	38	27	34	40	46	32	14	4	1	0	42	36	44	51	45	33	7	0	0	0	47	31	37	32	43	44	24	5	3	0
Buy	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sell	46	1	3	0	0	0	0	0	0	0	56	1	0	0	0	0	0	0	0	0	43	0	0	1	0	0	0	0	0	0
BA1*	0	2	2	3	14	19	17	1	0	0	5	1	7	17	11	0	0	0	0	0	0	0	10	6	9	13	18	20	9	7
AA1*	11	9	6	7	2	1	1	0	0	0	14	2	2	12	16	14	0	0	0	0	7	13	11	10	8	4	2	0	0	0
BC1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	0	0
BC2	1	0	0	0	20	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	1	0	0	0	18	0	0	0	0	0
BC3	0	3	0	0	0	31	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	27	0	0	0	0
BC4	56	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	51	0	0	0	0	1	5	0	0	0
BC5	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	1	0	0	0	0	0	0
BC6	0	0	3	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0
BC7	0	0	1	4	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	5	0	2	3	2	0	0	1	0	0
BC8	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
BC9	3	0	2	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
BC10	0	5	1	9	7	0	0	1	0	2	0	2	9	1	0	0	0	0	0	0	0	3	1	3	3	2	3	0	0	1
BC11	1	1	3	3	7	20	9	1	0	0	6	2	9	36	16	0	0	0	0	0	1	0	3	1	1	13	15	15	11	5
BC12	0	6	2	3	4	0	0	0	0	2	0	1	14	7	0	0	0	0	0	0	0	2	3	1	3	3	0	3	0	4
BC13	0	3	6	4	14	14	1	0	0	2	0	0	9	2	2	0	0	0	0	0	0	1	5	1	0	0	0	0	1	1
BC14	0	0	0	15	3	1	2	0	0	1	0	0	2	1	0	0	0	0	0	0	0	1	0	5	3	0	3	0	0	0
BC15	0	0	1	0	25	0	0	0	0	2	2	0	1	0	15	0	0	0	0	0	0	0	0	0	10	0	2	0	0	0
BC16	0	1	2	0	3	27	0	0	0	2	1	22	2	3	0	21	1	0	0	0	0	0	0	0	0	11	0	0	0	0
BC17	0	3	4	2	1	0	14	1	0	0	0	15	6	3	0	0	7	0	0	1	0	0	0	0	1	1	11	0	0	1
BC18	0	1	14	5	4	3	0	7	1	4	0	0	1	1	0	0	0	2	0	0	0	1	6	6	1	2	0	3	1	1
BC19	0	1	1	17	5	0	3	0	2	0	0	0	0	0	0	0	0	0	1	0	0	2	4	11	4	3	3	4	1	2
BC20	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
AC1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AC2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AC3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AC4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AC5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AC6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AC7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AC8	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
AC9	1	0	1	0	0	0	0	0	0	0	1	2	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0
AC10	3	2	6	2	0	1	0	0	0	0	3	3	3	1	3	3	1	0	0	0	2	6	1	2	1	0	1	1	0	0
AC11	1	6	6	11	1	0	0	0	0	0	11	1	0	8	8	6	0	0	0	0	7	4	1	6	4	1	0	0	0	0
AC12	4	5	3	6	3	2	3	2	0	0	0	0	0	0	0	0	0	1	0	0	1	2	6	1	1	1	3	1	0	0
AC13	0	2	0	1	0	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	1	1	3	0	0	0	0
AC14	1	2	1	1	2	0	0	2	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	2	0	0	0	0	0	1
AC15	2	1	2	1	0	0	0	1	1	0	1	2	0	1	1	2	1	0	0	0	0	0	0	0	1	1	0	0	0	0
AC16	2	4	4	3	0	0	3	1	0	0	2	1	1	1	4	2	0	1	0	0	1	3	1	2	0	0	0	0	0	0
AC17	5	6	3	3	2	2	1	0	0	0	3	2	0	1	1	1	1	0	1	1	0	1	3	4	2	1	0	0	0	0
AC18	5	2	1	4	1	0	0	1	0	0	3	1	1	3	0	0	0	0	0	0	0	1	1	0	1	1	0	0	1	1
AC19	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	2	2	1	0	0	0	0	0
AC20	0	1	3	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0

Note: Due to the general absence of relationships involving these explanatory events, BA11 to BA20 and AA11 to AA20 are omitted.

Panel IV: BC1 to BC10

	BMW										SAP										ADS									
BC	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Buy*	15	34	44	48	49	33	21	22	33	52	19	27	40	51	57	34	31	16	29	52	13	42	45	52	51	32	16	25	34	47
Sell*	61	61	61	60	60	61	60	50	46	13	61	61	61	59	58	59	54	33	17	29	61	61	61	61	61	59	58	47	35	21
Buy	12	19	21	15	13	5	2	1	0	1	4	11	7	21	8	4	3	2	0	0	9	15	15	7	10	11	0	0	0	0
Sell	59	58	61	57	49	45	31	36	23	0	59	61	58	53	44	30	20	14	0	0	58	56	53	40	29	33	19	13	8	0
BA1*	23	61	61	61	61	61	61	60	61	61	12	61	58	61	58	61	61	59	58	61	9	58	59	61	61	61	60	57	60	60
AA1*	61	61	61	60	59	44	13	5	4	1	61	60	54	51	42	27	23	6	1	2	61	61	61	60	57	54	31	5	2	0
BA1	57	55	27	40	8	1	1	0	0	0	61	36	32	27	2	0	0	0	0	0	52	33	4	5	2	0	0	0	0	0
BA2	0	38	2	11	16	10	15	0	0	0	0	55	6	17	8	4	14	0	0	0	4	24	4	3	1	4	22	7	0	1
BA3	0	2	40	0	1	1	0	4	0	0	0	5	55	6	8	1	0	0	0	0	0	3	26	4	1	2	2	22	8	0
BA4				29	0	2	0	3	4	0				59	2	0	0	0	0	0				10	2	1	0	0	12	0
BA5				0	41	1	0	0	7	4				1	59	0	0	0	0	0				2	28	0	0	1	10	12
BA6				0	0	25	1	0	0	3				0	0	57	0	1	0	0				1	1	16	1	0	1	4
BA7							2	3	0	1							2	8	1	1							3	5	1	4
BA8							1	2	0	0							0	4	4	0							0	0	2	1
BA9							0	0	2	0							2	3	4	0							0	0	4	0
BA10							2	0	0	1							1	0	0	5							1	0	0	3
AA1	60	61	61	29	3	0	0	0	0	0	61	60	42	8	1	0	0	0	0	0	61	60	59	25	10	4	1	0	0	0
AA2	52	54	52	21	1	1	0	0	0	0	53	55	22	4	0	0	0	0	0	0	51	49	49	22	7	2	0	0	0	0
AA3	35	51	54	6	0	0	0	0	0	0	27	49	20	2	1	0	0	0	0	0	40	52	52	19	4	3	0	0	0	0
AA4				1	0	0	0	0	0	0				0	1	0	0	0	0	0				6	0	1	1	0	0	0
AA5				4	0	0	0	0	0	0				0	0	0	0	0	0	0				0	2	0	0	0	0	0
AA6				0	1	0	0	0	0	0				2	2	0	0	0	0	0				0	0	0	0	0	0	0
AA7							0	0	0	0							0	1	0	0							0	0	0	0
AA8							0	0	1	0							1	0	0	0							0	0	0	0
AA9							1	0	0	0							1	0	0	0							0	0	0	0
AA10							0	0	0	0							0	1	0	0							0	0	0	0
BC1	61	61	61	60	49	2	2	0	0	0	61	61	61	50	27	5	0	0	0	0	61	61	61	61	58	25	5	1	1	0
BC2	61	61	61	36	17	0	0	0	0	0	61	61	60	43	14	1	0	0	0	0	61	60	61	37	41	18	1	0	0	0
BC3	59	61	60	60	46	36	8	0	0	0	43	60	50	45	46	3	0	0	0	0	55	61	60	49	48	44	23	2	0	0
BC4				52	48	43	18	3	0	1				6	25	20	0	0	0	0				35	46	46	31	8	1	0
BC5				47	19	43	15	14	0	0				37	2	19	2	0	0	0				52	22	50	34	11	4	0
BC6				23	36	25	43	10	2	1				6	17	9	28	1	0	0				25	30	2	32	26	3	0
BC7							33	16	5	4							19	22	6	0							16	30	23	8
BC8							21	7	6	3							8	1	9	5							25	1	15	9
BC9							2	13	1	4							5	14	3	5							1	12	1	11
BC10							19	14	28	6							8	11	14	2							8	5	26	3
AC1	5	32	29	44	15	2	4	0	1	6	0	11	24	39	11	1	0	1	0	1	3	16	8	33	5	0	0	0	0	3
AC2	0	4	6	0	0	0	1	0	0	0	0	3	11	7	5	2	0	0	0	0	2	4	3	0	1	0	0	0	0	0
AC3	1	20	39	16	6	2	0	0	0	0	0	15	32	16	3	0	0	0	0	0	7	8	12	3	5	4	2	0	0	3
AC4				16	16	2	0	0	0	2				28	16	1	0	0	0	0				4	1	0	0	1	0	1
AC5				6	4	1	0	0	0	0				14	12	3	0	0	0	0				5	3	1	0	0	0	2
AC6				3	4	2	0	0	0	0				5	10	1	0	0	0	0				4	2	1	0	0	0	0
AC7							0	1	0	0							0	0	0	0							1	0	0	0
AC8							0	0	0	0							0	1	1	1							1	0	0	1
AC9							0	0	0	0							1	1	1	3							0	0	1	4
AC10							0	6	3	0							4	9	2	3							3	1	3	4

Panel V: BC11 to BC20

	BMW										SAP										ADS									
BC	11	12	13	14	15	16	17	18	19	20	11	12	13	14	15	16	17	18	19	20	11	12	13	14	15	16	17	18	19	20
Buy*	54	51	44	36	41	40	29	30	15	1	42	39	35	43	37	49	38	23	3	0	57	50	38	28	42	44	34	31	12	10
Sell*	7	5	7	6	1	0	0	0	0	0	15	3	0	0	0	0	0	0	0	0	5	10	7	9	8	0	0	1	0	0
Buy	3	1	2	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
Sell	3	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
BA1*	60	59	51	36	34	39	53	46	33	0	61	50	29	30	42	40	42	17	0	0	61	59	52	40	38	41	50	52	49	15
AA1*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
BA1	2	0	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0
BA2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
BA4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
BA5	0	0	22	1	0	0	0	0	0	0	0	1	5	0	0	0	0	0	0	0	1	0	21	0	0	0	0	0	0	0
BA6	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	6	0	0	0	0	0	0
BA7	0	0	4	2	0	0	0	5	9	0	0	0	4	18	13	4	0	0	0	0	1	0	0	0	0	1	0	1	3	0
BA8	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA9	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
BA10	14	0	0	0	0	0	0	0	0	0	21	1	0	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0
AA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA9	0	0	0	0	0	0	0	0	0	1	0	0	2	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA10	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC11	2	1	4	11	16	9	12	16	8	0	0	7	28	14	12	20	16	1	0	0	6	2	2	7	3	12	8	15	10	2
BC12	4	2	1	7	8	4	5	5	3	3	1	2	0	15	7	2	2	0	0	0	5	1	6	4	7	5	4	8	4	2
BC13	1	2	4	0	7	14	13	4	0	3	1	0	2	3	8	0	0	0	0	0	3	1	1	1	3	1	3	1	2	0
BC14	1	4	6	4	1	2	1	5	0	0	1	3	4	0	2	1	1	0	1	0	0	1	1	2	4	2	5	0	2	0
BC15	1	2	2	2	12	1	2	2	1	0	1	3	6	2	7	9	4	0	0	0	1	3	1	3	4	2	4	1	3	0
BC16	6	3	1	4	3	8	2	4	1	1	1	3	11	3	4	16	3	2	2	0	0	0	2	2	3	7	3	1	2	0
BC17	4	3	5	5	3	2	10	4	3	4	1	4	5	6	5	2	14	1	0	0	0	3	1	4	1	1	16	5	2	0
BC18	1	1	4	9	0	6	4	12	7	1	1	2	2	2	3	7	3	1	0	0	1	1	2	4	4	1	2	7	4	1
BC19	1	3	5	2	3	0	7	6	2	0	0	0	1	0	0	0	0	2	0	0	0	1	5	4	1	10	3	10	5	2
BC20	0	0	2	4	4	2	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	2	2	0	1	1	1	2	0
AC11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
AC12	3	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	1	2	1	0	0	0	0	0
AC13	1	7	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	0	0	0	0
AC14	1	4	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
AC15	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	1	0	0
AC16	0	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
AC17	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1	0	0	1	0	0	0	0
AC18	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	0	0	0	0	0
AC19	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	3	0	0	1	0	0	0	0	0
AC20	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0

Note: Due to the general absence of relationships involving these explanatory events, BA11 to BA20, AA11 to AA20, BC1 to BC10 and AC1 to AC10 are omitted.

Panel VI: AA1 to AA10

	BMW										SAP										ADS									
AA	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Buy*	61	36	59	55	46	32	29	40	59	61	61	31	58	53	43	49	39	52	49	61	60	35	57	52	35	21	31	41	60	61
Sell*	61	61	60	58	49	21	14	11	0	1	61	60	51	48	50	31	21	10	2	0	61	61	61	59	54	31	18	16	4	1
Buy	15	12	18	10	6	6	5	2	1	5	39	13	5	4	10	4	4	3	0	11	23	4	14	3	5	3	1	0	1	9
Sell	58	59	52	29	9	1	0	1	0	0	51	53	26	4	0	0	0	0	0	0	50	54	41	15	6	2	2	3	0	0
BA1*	25	30	53	11	13	19	28	25	11	21	30	38	23	14	11	12	18	11	45	19	4	24	12	12	12	15	20	22	20	25
AA1*	61	61	61	61	56	55	23	4	0	1	53	61	61	60	44	31	26	5	2	9	60	61	61	61	52	48	34	9	2	1
BA1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
BA2	0	2	3	3	0	1	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	2	0	0	1	0
BA3	0	1	54	7	0	0	0	0	0	0	1	2	39	4	3	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0
BA4				18	1	0	1	0	0	0				16	2	0	0	0	0	0				17	0	0	0	0	0	0
BA5				9	12	2	0	0	0	1				2	1	0	0	0	0	0				1	2	1	0	0	0	0
BA6				15	4	1	1	0	0	0				7	2	1	0	0	0	0				3	0	4	0	0	0	0
BA7							1	0	0	0							3	0	0	0							1	0	1	0
BA8							0	2	0	0							0	0	0	0							0	0	0	0
BA9							3	0	0	1							2	0	2	4							1	1	0	0
BA10							0	0	0	0							0	0	0	12							0	0	0	1
AA1	61	61	61	59	7	0	0	0	0	0	59	61	58	21	11	0	0	0	0	8	51	61	61	45	20	5	1	1	0	0
AA2	54	61	59	19	5	0	0	0	0	0	39	18	41	18	6	1	0	0	0	0	25	57	57	13	4	3	1	0	0	0
AA3	31	60	59	37	6	0	2	0	0	0	6	59	43	48	9	2	1	1	0	0	20	48	56	34	25	14	11	1	0	0
AA4				29	51	2	1	0	0	0				7	38	2	1	0	1	0				10	37	12	3	0	0	0
AA5				44	19	44	3	0	3	5				0	0	25	0	1	3	2				30	7	47	4	0	0	3
AA6				12	28	6	42	11	3	14				10	1	0	27	5	1	3				3	30	2	46	0	0	6
AA7							0	50	9	7							2	26	4	3							1	55	2	2
AA8							8	1	56	4							5	2	37	2							8	1	57	3
AA9							6	30	1	59							2	13	1	52							3	26	1	58
AA10							6	3	0	0							3	3	0	1							6	4	1	0
BC1	61	61	61	59	28	0	1	0	0	0	61	60	61	43	11	1	0	0	0	0	61	61	60	61	49	4	0	0	0	0
BC2	53	57	57	32	0	0	0	0	0	0	41	51	51	3	1	0	0	0	0	0	22	59	57	32	8	0	0	0	0	0
BC3	57	60	60	49	4	0	0	0	0	0	24	39	48	8	3	0	0	0	0	0	32	51	59	40	8	0	0	0	0	0
BC4				34	4	0	0	0	0	0				4	8	3	0	0	0	0				33	7	10	4	0	0	0
BC5				13	0	0	0	1	0	0				5	7	7	2	0	0	0				16	8	5	10	1	0	0
BC6				27	0	0	0	2	0	0				12	4	3	4	1	0	0				18	0	0	3	1	0	0
BC7							2	5	2	0							1	4	0	0							1	3	5	0
BC8							1	0	2	2							1	4	0	1							0	0	0	1
BC9							2	0	0	0							0	3	0	1							1	0	1	0
BC10							10	10	0	0							6	13	6	1							3	3	0	0
AC1	61	59	61	28	21	17	16	14	1	59	61	43	58	32	14	21	7	13	0	57	61	59	53	29	27	15	22	39	14	58
AC2	25	39	60	35	2	1	20	11	21	0	32	42	58	34	10	6	13	2	50	0	41	55	61	47	5	5	27	37	58	0
AC3	14	60	59	61	26	4	12	8	0	59	10	35	61	59	42	2	2	2	0	4	24	51	31	61	17	6	10	11	1	60
AC4				46	51	18	4	1	1	0				61	44	48	1	1	0	0				34	54	13	4	0	1	1
AC5				53	57	9	4	1	10	0				41	61	27	6	2	1	0				58	57	31	6	1	12	0
AC6				58	9	57	0	0	0	0				49	41	61	14	3	0	1				58	31	58	19	1	0	5
AC7							46	6	0	0							39	7	1	0							47	11	3	0
AC8							1	43	14	5							12	31	0	0							5	42	14	2
AC9							0	4	21	3							2	4	7	1							1	2	16	2
AC10							3	2	5	11							1	2	2	1							6	0	4	10

Panel VII: AA11 to AA20

	BMW										SAP										ADS									
AA	11	12	13	14	15	16	17	18	19	20	11	12	13	14	15	16	17	18	19	20	11	12	13	14	15	16	17	18	19	20
Buy*	43	33	29	39	45	31	10	3	0	0	47	30	47	47	46	33	5	0	0	0	45	27	34	37	40	42	27	4	3	0
Sell*	0	0	2	8	3	2	2	0	0	8	1	1	25	12	1	0	0	0	0	0	0	1	0	1	0	4	1	2	2	13
Buy	4	1	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0
Sell	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA1*	6	8	7	12	2	1	0	0	0	0	12	5	2	12	16	8	0	0	0	0	4	12	9	8	7	10	3	0	0	0
AA1*	0	2	3	8	16	23	17	1	0	0	3	1	7	23	6	0	0	0	0	0	0	0	5	4	10	14	13	13	3	4
BC1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC8	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC9	1	1	0	1	0	1	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0
BC10	3	2	2	6	1	0	1	1	0	0	1	2	1	1	5	4	0	0	0	0	0	4	4	1	1	1	0	0	0	0
BC11	6	5	3	8	0	2	0	0	0	0	6	1	0	1	5	6	0	0	0	0	2	4	4	7	1	3	0	0	0	0
BC12	4	6	3	11	5	2	4	0	0	0	2	1	0	0	0	0	1	0	0	0	0	2	7	3	3	4	0	0	0	0
BC13	1	2	0	1	2	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0
BC14	0	2	2	2	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0
BC15	0	1	1	2	0	1	1	1	0	0	3	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1
BC16	1	5	2	3	3	1	1	4	0	0	6	1	1	2	3	3	1	0	2	0	1	1	1	4	0	0	0	1	0	0
BC17	2	5	4	8	2	0	3	0	0	1	2	3	0	1	5	2	2	0	0	0	0	2	1	2	0	1	1	1	0	0
BC18	3	6	3	2	1	0	0	2	0	0	0	0	0	1	1	2	0	0	0	0	0	0	0	3	2	0	0	0	0	0
BC19	4	3	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	3	1	0	1	0	1	0
BC20	2	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
AC1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0
AC2	0	0	0	0	17	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	24	0	0	0	0	0
AC3	0	1	0	0	1	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	3	30	1	0	0	0
AC4	56	0	0	0	0	1	7	0	0	0	0	0	0	0	0	0	0	0	0	0	54	0	0	0	0	0	6	0	0	0
AC5	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0
AC6	0	1	1	0	0	0	0	0	0	0	1	2	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
AC7	1	0	3	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	2	5	3	0	0	0	0	0
AC8	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	3	0	0	0	0	1	0	0	0
AC9	4	0	2	0	2	0	0	0	0	0	0	1	2	1	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
AC10	2	2	2	5	6	1	0	0	0	0	0	2	5	3	0	0	0	0	0	0	0	2	0	1	1	0	1	1	0	1
AC11	0	3	3	2	6	20	14	0	0	0	6	2	9	37	22	0	0	0	0	0	0	0	1	1	4	15	17	20	12	4
AC12	0	7	3	2	5	1	0	1	0	5	0	1	11	7	0	0	0	0	0	0	0	4	4	0	3	1	1	1	1	5
AC13	0	1	7	10	17	11	1	0	0	1	0	1	5	1	2	0	0	0	0	0	0	0	9	1	1	0	2	0	0	0
AC14	0	2	1	11	8	1	0	0	0	1	1	0	0	4	0	0	0	0	0	0	0	0	7	6	1	1	0	2	0	0
AC15	0	1	0	0	23	0	0	0	0	2	0	2	1	0	17	0	0	0	0	0	0	0	0	0	14	1	1	0	0	1
AC16	0	0	2	0	1	25	0	0	0	1	1	13	1	4	1	23	0	0	0	0	0	1	0	0	2	14	1	0	0	2
AC17	0	0	3	0	2	1	8	0	1	0	3	9	4	1	0	0	4	0	0	1	0	0	1	1	4	1	11	0	0	3
AC18	0	3	11	10	2	6	1	4	0	1	0	0	2	1	0	0	0	0	0	1	0	1	5	2	2	2	1	3	0	1
AC19	0	0	2	15	3	3	2	0	6	2	0	0	0	0	0	0	0	0	0	0	0	1	6	6	2	1	3	3	2	2
AC20	0	0	0	1	1	0	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0

Panel VIII: AC1 to AC10

	BMW										SAP										ADS									
AC	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Buy*	61	61	61	60	59	61	59	48	43	10	61	61	61	61	61	60	51	30	14	24	61	61	61	60	60	59	56	47	34	18
Sell*	11	35	47	54	43	33	21	25	31	51	26	37	40	51	57	31	27	23	25	53	16	51	50	48	46	32	23	29	32	51
Buy	59	61	58	50	47	44	23	30	16	0	61	58	59	47	40	32	15	9	0	0	60	53	54	37	25	33	23	19	6	1
Sell	13	28	18	18	15	3	4	1	1	0	10	14	11	14	10	2	0	0	0	0	4	12	9	5	7	9	0	0	0	0
BA1*	61	61	61	59	54	36	9	3	1	0	61	60	51	48	42	23	18	7	2	0	61	61	61	61	58	58	34	10	4	0
AA1*	17	61	60	61	60	61	61	60	61	61	10	61	59	59	57	61	61	57	56	60	7	59	59	61	61	61	61	59	59	61
BA1	61	61	61	37	3	0	0	0	0	0	61	60	50	10	1	0	0	0	0	0	61	60	60	29	11	4	0	0	0	0
BA2	51	57	50	15	1	0	0	0	0	0	50	40	22	4	1	0	0	0	0	0	52	51	49	20	10	4	0	0	0	0
BA3	26	53	54	3	0	0	0	0	0	0	18	39	34	3	0	0	0	0	0	0	35	45	56	22	9	0	0	0	0	0
BA4				1	0	0	0	0	0	0				3	0	0	0	0	0	0				6	0	0	0	0	0	0
BA5				2	1	0	0	0	0	0				1	2	0	0	0	0	0				2	0	0	0	0	0	0
BA6				1	0	0	0	0	0	0				7	5	2	0	0	0	0				0	0	0	0	0	0	0
BA7							0	0	0	0							0	1	0	0							0	0	0	0
BA8							0	0	1	1							0	1	1	0							0	0	0	0
BA9							1	0	0	0							1	1	2	0							0	0	0	0
BA10							0	0	0	0							0	0	0	0							0	0	0	0
AA1	54	56	21	37	9	2	1	0	0	0	60	28	19	22	15	3	0	0	0	0	48	26	1	4	4	1	0	0	0	0
AA2	0	39	2	6	14	5	11	1	0	0	0	46	5	5	6	4	16	2	0	0	1	26	2	2	1	4	27	12	0	3
AA3	1	0	41	1	0	2	0	3	0	0	0	1	55	10	1	2	1	6	0	1	0	0	29	4	3	6	4	25	5	0
AA4				16	0	2	0	8	6	0				59	1	0	0	0	0	0				10	0	0	0	2	8	0
AA5				1	37	2	0	0	10	4				3	60	0	0	0	1	0				1	29	0	0	1	7	5
AA6				0	0	27	0	0	0	3				0	3	46	0	0	1	0				2	0	11	1	0	1	3
AA7							2	4	0	2							4	6	1	0							2	9	1	3
AA8							0	3	2	0							0	3	3	0							0	1	2	0
AA9							0	0	3	0							4	1	4	0							0	0	1	1
AA10							0	0	0	5							1	0	1	6							0	0	0	0
BC1	2	35	30	48	16	1	5	1	0	2	1	19	26	34	10	2	1	2	1	1	7	9	5	24	12	5	0	0	0	6
BC2	0	10	6	1	0	0	0	0	0	0	0	11	14	9	10	0	0	0	0	1	3	1	4	1	3	0	0	0	0	0
BC3	4	18	39	11	5	5	0	1	0	3	3	16	27	20	9	1	0	0	0	0	7	16	17	5	5	3	0	1	0	0
BC4				18	15	0	0	0	0	0				31	21	6	0	0	0	0				5	8	1	0	0	0	0
BC5				13	9	1	0	0	0	1				18	16	1	0	0	0	0				7	3	2	0	0	1	1
BC6				5	4	1	0	0	1	0				15	14	8	0	0	0	0				5	3	0	0	0	0	0
BC7							0	0	0	0							1	1	0	0							0	0	0	2
BC8							0	0	1	1							0	1	2	1							0	0	1	1
BC9							0	0	1	3							1	2	2	3							0	0	1	2
BC10							2	1	2	1							8	8	4	3							1	4	4	8
AC1	61	61	61	61	45	3	4	0	0	0	61	61	60	54	29	5	0	0	0	0	61	61	61	59	58	26	6	4	1	0
AC2	60	61	61	33	12	1	0	0	0	0	61	61	60	43	13	4	0	0	0	0	60	60	61	38	37	12	1	1	0	0
AC3	54	61	61	60	51	31	3	0	0	0	31	60	55	44	46	7	4	2	0	0	48	60	61	54	46	44	10	0	0	0
AC4				45	47	46	27	2	0	0				7	32	20	1	1	1	0				38	46	47	37	7	0	3
AC5				44	22	45	21	6	0	1				18	2	21	7	3	5	0				43	18	53	38	14	0	0
AC6				26	39	25	47	21	1	0				6	20	8	25	6	1	0				21	24	2	37	24	5	0
AC7							32	20	3	3							18	22	7	1							7	27	17	4
AC8							22	2	9	9							8	4	11	2							29	1	17	10
AC9							5	12	1	3							2	6	1	5							3	8	0	6
AC10							21	17	31	1							3	1	13	6							8	11	17	2

Panel IX: AC11 to AC20

	BMW										SAP										ADS									
AC	11	12	13	14	15	16	17	18	19	20	11	12	13	14	15	16	17	18	19	20	11	12	13	14	15	16	17	18	19	20
Buy*	8	10	12	11	2	0	0	0	0	0	15	5	0	0	0	0	0	0	0	0	2	7	11	11	6	0	0	0	0	0
Sell*	52	55	45	39	42	49	35	28	15	6	44	48	33	39	31	46	43	24	5	0	60	49	40	37	35	48	38	28	13	8
Buy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0
Sell	4	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
BA1*	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA1*	61	59	42	30	29	32	54	48	35	0	60	42	32	22	41	34	42	21	2	0	61	58	49	37	40	41	50	57	48	11
BA1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BA9	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
BA10	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA1	1	1	1	0	0	0	0	0	0	0	10	0	1	0	0	0	0	0	0	0	5	0	0	0	0	1	0	0	0	0
AA2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
AA3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0
AA4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA5	0	0	17	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	1	0	23	0	0	0	0	0	0	0
AA6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0
AA7	0	0	2	0	0	0	0	2	7	0	0	0	5	22	15	6	0	0	0	0	0	0	0	0	0	0	0	6	9	0
AA8	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AA9	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
AA10	14	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	27	0	0	0	0	0	0	0	0	0
BC11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	3	3	0	0	0	0	0	0	0
BC13	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
BC14	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
BC15	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
BC16	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
BC17	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BC18	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0
BC19	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	1	0	0
BC20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AC11	1	1	2	10	8	12	10	9	5	0	3	9	25	17	23	24	5	6	0	0	2	1	7	8	8	10	13	9	10	3
AC12	6	0	2	6	4	9	4	7	3	0	2	0	1	11	7	1	0	2	2	0	2	0	1	3	4	4	8	5	6	2
AC13	1	3	2	1	4	8	9	4	2	0	2	3	3	4	1	3	1	0	0	0	2	2	0	1	3	3	4	1	3	1
AC14	1	1	1	2	4	4	5	6	3	1	1	5	1	1	3	1	0	1	0	0	1	1	4	2	0	2	3	2	3	2
AC15	4	2	0	2	8	2	7	0	0	0	3	2	5	2	9	8	5	1	0	0	1	2	2	5	5	1	2	1	0	0
AC16	6	1	2	0	1	3	2	2	3	0	2	3	8	4	4	15	3	1	0	0	1	1	2	1	0	6	2	2	1	0
AC17	0	5	2	8	4	2	10	5	4	2	1	0	6	1	10	5	10	1	0	0	1	2	3	3	0	1	10	6	4	1
AC18	2	3	4	15	3	2	12	9	10	0	1	2	5	3	2	0	3	0	0	0	1	3	1	5	2	5	5	11	2	1
AC19	4	1	2	3	3	4	5	9	5	0	1	1	1	1	0	0	0	2	1	0	1	0	1	4	4	6	2	6	6	0
AC20	2	2	1	4	0	2	2	3	0	0	0	1	1	0	0	0	0	0	0	0	1	0	3	3	0	0	0	2	2	1

Table 4.3 Descriptive models parameters - baselines

Panel I : Trades and LOS w/ BPI

	BMW						SAP						ADS					
m	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*
$\widehat{\mu_d^m}$	0.03	0.03	0.02	0.02	0.04	0.05	0.04	0.03	0.02	0.02	0.03	0.03	0.02	0.02	0.02	0.02	0.04	0.04
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.56	0.55	0.63	0.61	0.32	0.36	0.67	0.67	0.61	0.61	0.29	0.28	0.56	0.57	0.79	0.80	0.36	0.35

Panel II : LOS w/o BPI and LOC

	BMW																			
m	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11	BA12	BA13	BA14	BA15	BA16	BA17	BA18	BA19	BA20
$\widehat{\mu_d^m}$	0.10	0.12	0.12	0.08	0.07	0.04	0.06	0.03	0.03	0.05	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.00	0.01
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.23	0.34	0.24	0.20	0.33	0.35	0.61	0.37	0.40	0.37	0.47	1.00	0.78	0.76	0.78	0.56	1.00	1.00	1.00	1.00
m	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11	AA12	AA13	AA14	AA15	AA16	AA17	AA18	AA19	AA20
$\widehat{\mu_d^m}$	0.10	0.12	0.12	0.07	0.07	0.04	0.06	0.03	0.03	0.05	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.00	0.01
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.22	0.32	0.24	0.18	0.35	0.36	0.65	0.44	0.41	0.37	0.50	0.76	1.00	0.77	0.79	0.84	1.00	1.00	1.00	1.00
m	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11	BC12	BC13	BC14	BC15	BC16	BC17	BC18	BC19	BC20
$\widehat{\mu_d^m}$	0.08	0.09	0.08	0.08	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.00
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.17	0.18	0.20	0.28	0.23	0.27	0.36	0.56	0.46	0.44	0.26	0.54	0.54	0.51	0.51	0.51	0.52	0.42	0.42	1.00
m	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11	AC12	AC13	AC14	AC15	AC16	AC17	AC18	AC19	AC20
$\widehat{\mu_d^m}$	0.09	0.09	0.08	0.07	0.04	0.04	0.04	0.04	0.02	0.03	0.03	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.00	0.00
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.19	0.19	0.21	0.24	0.23	0.26	0.37	0.58	0.38	0.45	0.26	0.54	0.53	0.70	0.72	0.50	0.43	0.44	0.42	1.00
	SAP																			
m	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11	BA12	BA13	BA14	BA15	BA16	BA17	BA18	BA19	BA20
$\widehat{\mu_d^m}$	0.09	0.08	0.08	0.06	0.05	0.03	0.03	0.02	0.01	0.03	0.01	0.01	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.24	0.22	0.15	0.16	0.21	0.25	0.49	0.44	0.28	0.37	0.68	0.74	0.84	0.54	0.74	0.72	1.00	1.00	1.00	1.00
m	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11	AA12	AA13	AA14	AA15	AA16	AA17	AA18	AA19	AA20
$\widehat{\mu_d^m}$	0.09	0.08	0.07	0.07	0.06	0.03	0.04	0.02	0.01	0.03	0.01	0.01	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.24	0.22	0.14	0.16	0.22	0.29	0.60	0.52	0.26	0.39	0.79	1.00	0.84	0.54	0.74	0.76	1.00	1.00	1.00	1.00
m	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11	BC12	BC13	BC14	BC15	BC16	BC17	BC18	BC19	BC20
$\widehat{\mu_d^m}$	0.07	0.08	0.06	0.07	0.06	0.05	0.04	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.00
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.18	0.17	0.18	0.22	0.29	0.35	0.50	0.55	0.63	0.50	0.22	0.45	0.65	0.66	0.40	0.36	0.36	1.00	1.00	1.00
m	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11	AC12	AC13	AC14	AC15	AC16	AC17	AC18	AC19	AC20
$\widehat{\mu_d^m}$	0.07	0.08	0.06	0.07	0.06	0.05	0.05	0.04	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.00
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.18	0.17	0.17	0.21	0.27	0.34	0.56	0.65	0.63	0.50	0.21	0.44	0.41	0.67	0.42	0.33	0.35	1.00	1.00	1.00
	ADS																			
m	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11	BA12	BA13	BA14	BA15	BA16	BA17	BA18	BA19	BA20
$\widehat{\mu_d^m}$	0.08	0.06	0.08	0.05	0.04	0.04	0.04	0.02	0.02	0.03	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.32	0.33	0.30	0.22	0.31	0.32	0.52	0.33	0.35	0.37	0.48	0.75	0.75	0.73	0.75	0.78	1.00	1.00	1.00	1.00
m	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11	AA12	AA13	AA14	AA15	AA16	AA17	AA18	AA19	AA20
$\widehat{\mu_d^m}$	0.08	0.06	0.07	0.05	0.05	0.04	0.04	0.02	0.02	0.03	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.33	0.32	0.27	0.20	0.34	0.32	0.44	0.35	0.35	0.37	0.52	1.00	0.75	0.74	0.76	0.79	1.00	1.00	1.00	1.00
m	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11	BC12	BC13	BC14	BC15	BC16	BC17	BC18	BC19	BC20
$\widehat{\mu_d^m}$	0.05	0.06	0.06	0.06	0.05	0.05	0.03	0.04	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.19	0.23	0.28	0.33	0.42	0.41	0.44	0.64	0.55	0.51	0.25	0.51	0.55	0.61	0.52	0.49	0.48	0.42	0.46	1.00
m	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11	AC12	AC13	AC14	AC15	AC16	AC17	AC18	AC19	AC20
$\widehat{\mu_d^m}$	0.05	0.07	0.06	0.06	0.05	0.05	0.03	0.04	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00
$\widehat{\mu_d^m}/E[\lambda_d^m(t)]$	0.19	0.25	0.28	0.36	0.42	0.42	0.45	0.64	0.55	0.51	0.25	0.51	0.56	0.54	0.50	0.47	0.45	0.51	0.47	1.00

Table 4.4 Descriptive models parameters – Hawkes events effects

Panel I: Buy*, Sell*, Buy, Sell, BA1* and AA1*

		BMW						SAP						ADS					
n \ m		Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*
Buy*	$\overline{\hat{\alpha}_{n/m}}$					27	5.5					31	12					26	12
	$\overline{\hat{\beta}_{n/m}}$					101	29					118	86					90	92
Sell*	$\overline{\hat{\alpha}_{n/m}}$					8.9	28					11	35					13	27
	$\overline{\hat{\beta}_{n/m}}$					71	103					76	123					88	89
Buy	$\overline{\hat{\alpha}_{n/m}}$	38		8.5		10		19		8.8		11		34		4.9		14	
	$\overline{\hat{\beta}_{n/m}}$	162		45		87		107		41		108		169		27		113	
Sell	$\overline{\hat{\alpha}_{n/m}}$		36		9.0		11		20		7.8		11		33		4.1		14
	$\overline{\hat{\beta}_{n/m}}$		154		42		86		113		37		106		164		23		101
BA1*	$\overline{\hat{\alpha}_{n/m}}$	3.5	17		1.6	15			14		1.9			3.7	9.1			20	
	$\overline{\hat{\beta}_{n/m}}$	63	218		47	341			144		38			69	155			348	
AA1*	$\overline{\hat{\alpha}_{n/m}}$	17	3.4	1.5			15	11		1.9				8.5	3.6				19
	$\overline{\hat{\beta}_{n/m}}$	215	61	44			298	108		33				137	62				337
BA1	$\overline{\hat{\alpha}_{n/m}}$					11						12						8.5	
	$\overline{\hat{\beta}_{n/m}}$					232						269						198	
AA1	$\overline{\hat{\alpha}_{n/m}}$						8.2						12						8.5
	$\overline{\hat{\beta}_{n/m}}$						147						268						188
BC1	$\overline{\hat{\alpha}_{n/m}}$					2.0	14					0.6	10					4.7	12
	$\overline{\hat{\beta}_{n/m}}$					74	309					16	283					93	237
AC1	$\overline{\hat{\alpha}_{n/m}}$					14						11	0.8					12	5.1
	$\overline{\hat{\beta}_{n/m}}$					327						299	24					233	113

For each Hawkes effect n/m, this figure presents estimated parameters $\hat{\alpha}_d^{n/m}$ and $\hat{\beta}_d^{n/m}$ daily average.

Panel II: BA1 to BA10

		BMW										SAP										ADS									
n \ m		BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10
Buy*	$\hat{\alpha}_{n/m}$	95	78	56	35	17						89	69	46	21	9.5						68	69	55	43	21					
	$\hat{\beta}_{n/m}$	212	233	183	158	94						201	194	146	137	95						170	200	148	182	115					
Sell*	$\hat{\alpha}_{n/m}$	111	6.7	89	11	11	5.7		7.2	4.6	88	129	10	103	12	14	7.5	5.6	7.6	10	68	65	7.8	80	11	15			8.8	10	83
	$\hat{\beta}_{n/m}$	348	132	313	69	63	26		29	23	333	328	190	263	52	81	57	33	39	109	242	268	132	286	88	67			68	63	281
Buy	$\hat{\alpha}_{n/m}$	45	60	32								28	37	14								54	57	29							
	$\hat{\beta}_{n/m}$	207	475	369								184	401	240								273	434	293							
Sell	$\hat{\alpha}_{n/m}$											3.8																			
	$\hat{\beta}_{n/m}$											60																			
BA1*	$\hat{\alpha}_{n/m}$	39	65	101	41	21	9.3					19	48	115	17	4.7	2.8	4.2				28	40	97	43	22	9.0	5.9			
	$\hat{\beta}_{n/m}$	269	393	322	179	130	98					197	287	231	76	39	36	76				217	345	320	195	185	80	71			
AA1*	$\hat{\alpha}_{n/m}$			10																4.5											
	$\hat{\beta}_{n/m}$			316																171											
BA1	$\hat{\alpha}_{n/m}$	26	34	52	19							16	39	24								16	22	24	11						
	$\hat{\beta}_{n/m}$	353	604	770	506							240	593	463								247	486	434	426						
BA2	$\hat{\alpha}_{n/m}$	3.0	6.0	8.6								2.5		13									9.0	5.9							
	$\hat{\beta}_{n/m}$	53	106	136								64		262									216	124							
BA3	$\hat{\alpha}_{n/m}$		5.5	7.9	2.2								13	14	6.4								8.1	8.2	3.5						
	$\hat{\beta}_{n/m}$		105	107	50								308	209	163								266	177	99						
BA4	$\hat{\alpha}_{n/m}$					3.7										2.8										2.3					
	$\hat{\beta}_{n/m}$					71										91										70					
BA5	$\hat{\alpha}_{n/m}$				9.3		3.6																				3.4				
	$\hat{\beta}_{n/m}$				221		58																				62				
BA6	$\hat{\alpha}_{n/m}$							3.3										1.5								13		2.3			
	$\hat{\beta}_{n/m}$							29										39								344		30			
BA7	$\hat{\alpha}_{n/m}$								3.3										1.9										3.0		
	$\hat{\beta}_{n/m}$								34										30										49		
BA8	$\hat{\alpha}_{n/m}$									5.8										3.4										5.0	
	$\hat{\beta}_{n/m}$									41										109										42	
BA9	$\hat{\alpha}_{n/m}$								6.6		18										12										16
	$\hat{\beta}_{n/m}$								109		56										49										59
AA3	$\hat{\alpha}_{n/m}$			2.6																											
	$\hat{\beta}_{n/m}$			83																											
BC1	$\hat{\alpha}_{n/m}$	61	24	34							28	44	15	33	9.5		3.3				24	46	23	39	6.6				8.2		35
	$\hat{\beta}_{n/m}$	626	769	531							572	496	564	461	341		117				560	442	555	517	249				244		538
BC2	$\hat{\alpha}_{n/m}$	1.4	12	61	2.2							0.8	2.2	39	13					19		2.9	13	54	2.0				2.7	17	
	$\hat{\beta}_{n/m}$	50	516	975	57							17	50	825	370					681		63	447	884	47				75	362	
BC3	$\hat{\alpha}_{n/m}$		31	5.7	45						39		14	0.9	10	8.1							20		14						31
	$\hat{\beta}_{n/m}$		865	135	582						840		514	6.2	165	177							592		139						597
BC4	$\hat{\alpha}_{n/m}$				8.5	11									0.5	4.1	2.7									13					
	$\hat{\beta}_{n/m}$				229	174									1.6	96	77									235					
BC5	$\hat{\alpha}_{n/m}$				67	4.0										0.8									68	2.8	4.9				
	$\hat{\beta}_{n/m}$				897	40										2.2									875	35	76				
BC6	$\hat{\alpha}_{n/m}$				4.0		2.8								2.4		1.4								3.0		1.3				
	$\hat{\beta}_{n/m}$				67		36								24		6.7								16		11				
BC7	$\hat{\alpha}_{n/m}$							4.1										1.3										2.3			
	$\hat{\beta}_{n/m}$							38										18										19			
BC8	$\hat{\alpha}_{n/m}$								3.3										1.4										2.7		
	$\hat{\beta}_{n/m}$								45										33										46		
AC1	$\hat{\alpha}_{n/m}$	42	31	41	24							30	27	28	8.1							27	16	26	19	7.1					
	$\hat{\beta}_{n/m}$	458	552	561	491							356	401	427	248							283	328	333	375	166					
AC2	$\hat{\alpha}_{n/m}$	5.2	11	22								2.6	6.2	10								3.2	6.9	12	6.9						
	$\hat{\beta}_{n/m}$	127	313	507								63	131	211								74	200	264	234						
AC3	$\hat{\alpha}_{n/m}$	4.1	11	18	7.5								2.9	3.3								1.8	2.8	10	5.3						
	$\hat{\beta}_{n/m}$	89	220	205	196								64	57								43	81	182	157						
AC4	$\hat{\alpha}_{n/m}$																								3.6						
	$\hat{\beta}_{n/m}$																								123						

Panel III: AA1 to AA10

		BMW										SAP										ADS									
n \ m		AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10
Buy*	$\hat{\alpha}_{n/m}$	115	8.0	96	11	13	6.7		7.9	5.2	95	139	8.8	94	9.1	17	7.9	5.0	5.8	12	86	67	10	88	11	14		7.4	10	8.5	94
	$\hat{\beta}_{n/m}$	338	183	315	79	71	30		29	26	352	332	185	260	42	110	46	26	32	121	269	274	168	322	86	66		34	78	49	308
Sell*	$\hat{\alpha}_{n/m}$	88	76	54	31	16						78	61	44	21	11	4.7					64	64	55	42	19	6.6				
	$\hat{\beta}_{n/m}$	193	223	174	162	89						174	166	135	125	94	91					149	181	138	174	107	52				
Buy	$\hat{\alpha}_{n/m}$											3.9																			
	$\hat{\beta}_{n/m}$											66																			
Sell	$\hat{\alpha}_{n/m}$	44	55	29								24	35									55	54	27							
	$\hat{\beta}_{n/m}$	203	436	333								159	366									272	405	281							
BA1*	$\hat{\alpha}_{n/m}$			10									14							5.0											
	$\hat{\beta}_{n/m}$			285									263							174											
AA1*	$\hat{\alpha}_{n/m}$	36	61	107	37	18	8.6					14	41	99	19	4.6	2.2					21	40	94	43	19	7.4	4.8			
	$\hat{\beta}_{n/m}$	253	344	321	169	122	101					110	221	196	80	38	42					163	330	320	199	180	75	73			
BA3	$\hat{\alpha}_{n/m}$			2.7										3.0																	
	$\hat{\beta}_{n/m}$			81										117																	
AA1	$\hat{\alpha}_{n/m}$	25	33	52	17							16	36	24								12	19	22	10						
	$\hat{\beta}_{n/m}$	332	559	763	530							244	528	484								172	416	398	421						
AA2	$\hat{\alpha}_{n/m}$	3.3	5.5	8.4								2.5		9.1									10	6.0							
	$\hat{\beta}_{n/m}$	61	91	130								56		179									237	133							
AA3	$\hat{\alpha}_{n/m}$	1.1	5.4	7.1	2.5								14	13	5.1								7.5	10	2.9						
	$\hat{\beta}_{n/m}$	30	106	91	61								338	203	151								212	231	78						
AA4	$\hat{\alpha}_{n/m}$					4.6										3.5										2.9					
	$\hat{\beta}_{n/m}$					80										112										81					
AA5	$\hat{\alpha}_{n/m}$				12		3.5																				3.7				
	$\hat{\beta}_{n/m}$				302		53																				64				
AA6	$\hat{\alpha}_{n/m}$							3.6																				2.5			
	$\hat{\beta}_{n/m}$							34																				51			
AA7	$\hat{\alpha}_{n/m}$								3.5																				3.2		
	$\hat{\beta}_{n/m}$								41																				52		
AA8	$\hat{\alpha}_{n/m}$									6.5										3.9										5.6	
	$\hat{\beta}_{n/m}$									45										43										48	
AA9	$\hat{\alpha}_{n/m}$										19										11										14
	$\hat{\beta}_{n/m}$										57										55										56
BC1	$\hat{\alpha}_{n/m}$	41	34	41	23							27	27	29	6.9							24	17	28	18	6.6					
	$\hat{\beta}_{n/m}$	448	565	554	559							328	419	443	208							245	328	345	366	151					
BC2	$\hat{\alpha}_{n/m}$	5.3	11	23	6.9							2.6	6.0	9.2									7.5	12	6.7						
	$\hat{\beta}_{n/m}$	136	274	516	290							61	125	202									205	264	234						
BC3	$\hat{\alpha}_{n/m}$	4.3	11	17	8.4								2.2	3.1								1.2	3.6	10	5.8						
	$\hat{\beta}_{n/m}$	96	200	191	280								46	55								24	102	178	170						
BC4	$\hat{\alpha}_{n/m}$				4.0																				3.5						
	$\hat{\beta}_{n/m}$				135																				131						
AC1	$\hat{\alpha}_{n/m}$	59	24	35							29	48	14	30	6.0						26	47	25	45					8.5		37
	$\hat{\beta}_{n/m}$	575	761	530							585	542	585	466	170						551	492	593	558					278		547
AC2	$\hat{\alpha}_{n/m}$		13	60	1.5							0.8	3.3	43	10					20		3.1	14	56	3.7				2.6	19	
	$\hat{\beta}_{n/m}$		511	979	31							20	83	903	344					659		72	479	946	98				79	410	
AC3	$\hat{\alpha}_{n/m}$		33	6.2	46						38		14	1.4	15	6.6							19	5.4	14						34
	$\hat{\beta}_{n/m}$		864	131	620						831		525	8.4	268	175							610	215	127						649
AC4	$\hat{\alpha}_{n/m}$				10	11									0.8	3.6	3.6								8.2	14					
	$\hat{\beta}_{n/m}$				279	163									2.2	78	72								268	251					
AC5	$\hat{\alpha}_{n/m}$				60	3.2									16	2.5									65	4.1	4.9				
	$\hat{\beta}_{n/m}$				914	32									311	7.6									896	63	86				
AC6	$\hat{\alpha}_{n/m}$				4.1		3.3								2.6	8.8	4.7								2.9	11	1.7				
	$\hat{\beta}_{n/m}$				76		45								24	124	21								18	240	18				
AC7	$\hat{\alpha}_{n/m}$							4.4										1.5										2.5			
	$\hat{\beta}_{n/m}$							43										19										29			
AC8	$\hat{\alpha}_{n/m}$								3.4										1.5										4.3		
	$\hat{\beta}_{n/m}$								51										33										89		

Panel IV: BC1 to BC10

		BMW										SAP										ADS									
n \ m		BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10
Buy*	$\hat{\alpha}_{n/m}$		10	18	24	14	7.5			4.2	5.0			16	21	13	5.0	6.1			8.2		10	10	16	10	8.1			4.8	6.0
	$\hat{\beta}_{n/m}$		94	145	169	105	77			33	34			164	144	88	36	47			57		70	64	110	68	76			38	39
Sell*	$\hat{\alpha}_{n/m}$	119	167	78	39	24	19	14	8.2	5.1		108	181	60	36	22	16	10	2.8			113	151	69	33	22	17	15	8.2	5.6	
	$\hat{\beta}_{n/m}$	149	199	142	158	124	185	114	78	81		125	193	152	164	132	112	110	61			153	185	143	123	102	116	161	71	85	
Buy	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Sell	$\hat{\alpha}_{n/m}$	16	63	51	32	15	9.3	9.2	11			13	63	36	17	8.9						25	54	43	18		7.2				
	$\hat{\beta}_{n/m}$	84	403	330	246	183	138	134	167			87	357	289	239	150						145	364	291	206		88				
BA1*	$\hat{\alpha}_{n/m}$		57	30	128	40	10	12	10	15	39		61	24	138	21	8.1	6.5	7.7	6.5	18		15	20	103	24	8.3	9.3	6.9	10	27
	$\hat{\beta}_{n/m}$		250	298	496	327	27	43	56	115	236		255	250	444	142	24	28	55	68	207		87	187	450	258	37	48	50	93	187
AA1*	$\hat{\alpha}_{n/m}$	121	63	53	25	11	5.6					116	59	28	11	6.3						72	45	47	20	8.4	5.9	3.6			
	$\hat{\beta}_{n/m}$	421	360	356	242	156	145					383	262	223	149	107						270	243	277	202	113	98	91			
BA1	$\hat{\alpha}_{n/m}$	5.2	26		18							3.4	16	19								4.8	10								
	$\hat{\beta}_{n/m}$	53	891		648							18	619	708								66	324								
BA2	$\hat{\alpha}_{n/m}$		6.8										1.1																		
	$\hat{\beta}_{n/m}$		223										7.6																		
BA3	$\hat{\alpha}_{n/m}$			2.6										0.8																	
	$\hat{\beta}_{n/m}$			81										5.8																	
BA4	$\hat{\alpha}_{n/m}$														2.5																
	$\hat{\beta}_{n/m}$														46																
BA5	$\hat{\alpha}_{n/m}$					0.3										3.2															
	$\hat{\beta}_{n/m}$					2.9										32															
BA6	$\hat{\alpha}_{n/m}$																0.6														
	$\hat{\beta}_{n/m}$																2.7														
AA1	$\hat{\alpha}_{n/m}$	24	36	34								19	24	12								18	16	13							
	$\hat{\beta}_{n/m}$	392	682	687								289	414	370								271	333	320							
AA2	$\hat{\alpha}_{n/m}$	7.1	6.7	7.2								4.5	3.0									6.1	3.7	3.7							
	$\hat{\beta}_{n/m}$	205	199	235								115	65									154	111	101							
AA3	$\hat{\alpha}_{n/m}$	1.4	2.2	3.2									1.4									2.3	3.5	3.8							
	$\hat{\beta}_{n/m}$	39	65	91									26									64	96	96							
BC1	$\hat{\alpha}_{n/m}$	67	93	57	27	8.9						46	70	24	10							57	86	37	14	7.9					
	$\hat{\beta}_{n/m}$	391	460	552	593	260						303	393	376	317							348	462	382	297	194					
BC2	$\hat{\alpha}_{n/m}$	10	7.6	17	4.6							8.8	22	10	2.6							11	10	14	4.8	3.3					
	$\hat{\beta}_{n/m}$	133	80	284	158							158	290	190	76							157	149	213	151	110					
BC3	$\hat{\alpha}_{n/m}$	5.0	10	15	8.1	4.4	2.9					1.0	11	13	5.3	1.5						3.0	8.1	7.9	5.0	4.5	3.6				
	$\hat{\beta}_{n/m}$	87	125	192	154	121	92					25	213	244	147	26						62	137	110	102	122	96				
BC4	$\hat{\alpha}_{n/m}$				6.8	4.2	2.7																		3.6	3.1	3.5	2.7			
	$\hat{\beta}_{n/m}$				128	99	70																		76	67	90	74			
BC5	$\hat{\alpha}_{n/m}$				4.6		3.4								9.0										6.7		4.6	4.0			
	$\hat{\beta}_{n/m}$				70		81								296										147		103	114			
BC6	$\hat{\alpha}_{n/m}$					2.4		4.7																				3.8			
	$\hat{\beta}_{n/m}$					55		89																				115			
BC7	$\hat{\alpha}_{n/m}$							1.7																							
	$\hat{\beta}_{n/m}$							19																							
BC8	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
BC9	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
BC10	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
AC1	$\hat{\alpha}_{n/m}$		10		21										13										10						
	$\hat{\beta}_{n/m}$		420		569										387										329						
AC2	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
AC3	$\hat{\alpha}_{n/m}$			7.3										2.2																	
	$\hat{\beta}_{n/m}$			217										61																	
AC4	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														

Panel V: AC1 to AC10

		BMW										SAP										ADS									
n \ m		AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10
Buy*	$\hat{\alpha}_{n/m}$	117	170	72	36	22	19	13	8.8	5.2		104	170	62	34	23	17	10				119	153	71	31	21	16	15	8.7	5.8	
	$\hat{\beta}_{n/m}$	146	200	139	161	126	182	110	83	88		126	185	164	161	157	124	105				162	194	156	126	106	123	169	76	92	
Sell*	$\hat{\alpha}_{n/m}$		10	18	21	12	7.7			7.1	5.5		13	20	22	14	5.2				8.5		16	11	16	12	9.5			5.6	5.9
	$\hat{\beta}_{n/m}$		98	141	150	79	77			65	37		192	138	144	97	40				56		101	68	82	73	82			45	38
Buy	$\hat{\alpha}_{n/m}$	17	63	50	28	13	8.1					11	61	33	13	7.8	5.5					19	51	37	17		5.3				
	$\hat{\beta}_{n/m}$	87	422	347	259	178	139					69	369	310	214	154	98					97	386	276	216		75				
Sell	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
BA1*	$\hat{\alpha}_{n/m}$	117	69	57	24	10	5.3					112	63	27	12	11						67	48	48	20	8.3	6.0	3.5			
	$\hat{\beta}_{n/m}$	397	380	368	262	150	133					386	294	232	172	118						267	256	283	210	114	96	83			
AA1*	$\hat{\alpha}_{n/m}$		55	26	130	35	9.0	11	9.3	14	36		61	20	118	14	8.3	7.0	6.6	5.0	16		11	17	99	21	7.7	7.9	6.7	9.1	27
	$\hat{\beta}_{n/m}$		234	263	495	304	25	40	50	124	228		244	196	425	98	22	24	43	49	179		70	156	395	215	36	42	46	84	190
BA1	$\hat{\alpha}_{n/m}$	22	39	34	13							19	28	12								17	17	14							
	$\hat{\beta}_{n/m}$	313	683	694	497							280	467	366								278	361	342							
BA2	$\hat{\alpha}_{n/m}$	6.9	7.5	8.5								5.5	3.4									5.4	4.2	4.2							
	$\hat{\beta}_{n/m}$	179	234	291								168	78									153	115	125							
BA3	$\hat{\alpha}_{n/m}$		2.1	2.9									2.1	1.7								2.1	3.5	4.5							
	$\hat{\beta}_{n/m}$		56	82									47	50								57	89	113							
AA1	$\hat{\alpha}_{n/m}$	5.2	26		14							3.1										3.8									
	$\hat{\beta}_{n/m}$	57	856		515							20										52									
AA2	$\hat{\alpha}_{n/m}$		9.3										1.0																		
	$\hat{\beta}_{n/m}$		279										6.8																		
AA3	$\hat{\alpha}_{n/m}$			3.9										1.1																	
	$\hat{\beta}_{n/m}$			105										7.6																	
AA4	$\hat{\alpha}_{n/m}$														2.5																
	$\hat{\beta}_{n/m}$														44																
AA5	$\hat{\alpha}_{n/m}$					0.3										2.7															
	$\hat{\beta}_{n/m}$					2.7										48															
AA6	$\hat{\alpha}_{n/m}$																0.5														
	$\hat{\beta}_{n/m}$																1.6														
BC1	$\hat{\alpha}_{n/m}$		11		20										13																
	$\hat{\beta}_{n/m}$		411		607										373																
BC2	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
BC3	$\hat{\alpha}_{n/m}$			7.3																											
	$\hat{\beta}_{n/m}$			219																											
BC4	$\hat{\alpha}_{n/m}$														1.0																
	$\hat{\beta}_{n/m}$														42																
AC1	$\hat{\alpha}_{n/m}$	64	93	56	26	8.9						41	68	24	6.0							58	89	38	14	7.8					
	$\hat{\beta}_{n/m}$	354	455	572	676	275						272	365	364	173							352	465	397	310	192					
AC2	$\hat{\alpha}_{n/m}$	8.0	7.3	16	4.2							7.9	14	11	3.5							8.8	9.1	14	5.0	3.5					
	$\hat{\beta}_{n/m}$	95	79	282	160							127	166	232	97							134	119	212	153	117					
AC3	$\hat{\alpha}_{n/m}$	3.4	9.3	16	8.3	4.6	3.0					1.7	15	13	6.2	1.5						2.6	8.1	7.3	5.8	4.0	3.5				
	$\hat{\beta}_{n/m}$	51	117	185	160	125	100					49	254	230	170	27						53	136	98	117	109	98				
AC4	$\hat{\alpha}_{n/m}$				5.7	4.2	3.1									3.2									4.0	3.1	3.7	3.0			
	$\hat{\beta}_{n/m}$				101	96	80									106									73	68	93	82			
AC5	$\hat{\alpha}_{n/m}$				3.7		3.8																		5.6		4.9	4.2			
	$\hat{\beta}_{n/m}$				58		87																		122		105	122			
AC6	$\hat{\alpha}_{n/m}$					1.7		5.9																				5.0			
	$\hat{\beta}_{n/m}$					34		100																				143			
AC7	$\hat{\alpha}_{n/m}$							1.3																							
	$\hat{\beta}_{n/m}$							15																							
AC8	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
AC9	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
AC10	$\hat{\alpha}_{n/m}$									1.4																					
	$\hat{\beta}_{n/m}$									15																					

Panel VI: BA11 to BA20, AA11 to AA20, BC11 to BC20 and AC11 to AC20

		BMW										SAP										ADS									
n	m	BA11	BA12	BA13	BA14	BA15	BA16	BA17	BA18	BA19	BA20	BA11	BA12	BA13	BA14	BA15	BA16	BA17	BA18	BA19	BA20	BA11	BA12	BA13	BA14	BA15	BA16	BA17	BA18	BA19	BA20
Buy*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Sell*	$\hat{\alpha}_{n/m}$	5.4		1.4	3.3	7.6	3.9					3.3	0.8	2.6	4.7	7.0	3.1					6.3	3.4	1.8	3.0	6.3	6.8				
	$\hat{\beta}_{n/m}$	66		26	41	67	88					52	20	51	64	74	67					62	39	26	35	56	63				
Buy	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Sell	$\hat{\alpha}_{n/m}$	24										25										24									
	$\hat{\beta}_{n/m}$	351										413										310									
BA1*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
AA1*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
n	m	AA11	AA12	AA13	AA14	AA15	AA16	AA17	AA18	AA19	AA20	AA11	AA12	AA13	AA14	AA15	AA16	AA17	AA18	AA19	AA20	AA11	AA12	AA13	AA14	AA15	AA16	AA17	AA18	AA19	AA20
Buy*	$\hat{\alpha}_{n/m}$	5.2	2.3		3.2	7.3	5.7					3.4		2.6	4.5	6.7	3.3					5.9		1.8	3.3	6.3	6.9				
	$\hat{\beta}_{n/m}$	66	35		43	68	73					56		54	68	76	79					65		29	38	58	66				
Sell*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Buy	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Sell	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
BA1*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
AA1*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
n	m	BC11	BC12	BC13	BC14	BC15	BC16	BC17	BC18	BC19	BC20	BC11	BC12	BC13	BC14	BC15	BC16	BC17	BC18	BC19	BC20	BC11	BC12	BC13	BC14	BC15	BC16	BC17	BC18	BC19	BC20
Buy*	$\hat{\alpha}_{n/m}$	7.5	7.1	4.0	3.9	5.2	5.8					2.6	3.1	1.8	2.5	2.7	5.5	3.9				8.3	7.9	4.0		4.4	5.1	4.6	3.8		
	$\hat{\beta}_{n/m}$	40	54	42	50	62	60					23	42	17	23	40	83	86				43	61	43		63	59	51	53		
Sell*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Buy	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Sell	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
BA1*	$\hat{\alpha}_{n/m}$	362	12	4.8	2.5	2.3	3.8	4.8	2.2	0.6		322	3.0			3.1	2.3	2.8				324	13	4.8	2.7	2.0	2.8	4.7	3.8	2.5	
	$\hat{\beta}_{n/m}$	578	148	80	44	55	72	55	28	15		532	57			50	35	82				513	148	86	44	55	66	84	65	39	
AA1*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
n	m	AC11	AC12	AC13	AC14	AC15	AC16	AC17	AC18	AC19	AC20	AC11	AC12	AC13	AC14	AC15	AC16	AC17	AC18	AC19	AC20	AC11	AC12	AC13	AC14	AC15	AC16	AC17	AC18	AC19	AC20
Buy*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Sell*	$\hat{\alpha}_{n/m}$	7.4	7.8	4.5	4.3	4.7	6.0	3.9				2.5	3.7	2.1	2.8	3.0	6.5	3.9				8.7	7.9	3.9	3.8	4.1	5.7	4.9			
	$\hat{\beta}_{n/m}$	40	61	46	32	39	63	54				21	47	31	25	40	83	76				44	59	41	50	55	59	54			
Buy	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
Sell	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
BA1*	$\hat{\alpha}_{n/m}$																														
	$\hat{\beta}_{n/m}$																														
AA1*	$\hat{\alpha}_{n/m}$	353	11	3.8			3.4	4.5	2.3	0.6		317	2.8	1.3		3.3	2.1	2.5				323	11	4.6	2.0	2.3	2.6	4.7	4.4	2.6	
	$\hat{\beta}_{n/m}$	570	137	64			69	66	31	16		510	59	26		50	29	73				510	141	86	47	61	62	77	58	41	

Table 4.5 Depth Levels 1 to 11 descriptive models effects metrics

Panel I: Trades and LOS w/ BPI

	BMW						SAP						ADS					
	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*	Buy*	Sell*	Buy	Sell	BA1*	AA1*
Buy*					0.27 0.11	0.23 0.10					0.27 0.13	0.17 0.08					0.29 0.11	0.19 0.07
Sell*					0.18 0.07	0.28 0.12					0.17 0.08	0.28 0.14					0.18 0.07	0.31 0.11
Buy	0.22 0.12		0.20 0.20		0.13 0.03		0.17 0.13		0.22 0.22		0.10 0.04		0.20 0.11		0.21 0.21		0.13 0.03	
Sell		0.23 0.12		0.23 0.23		0.13 0.03		0.18 0.13		0.23 0.23		0.11 0.04		0.21 0.11		0.20 0.20		0.15 0.03
BA1*	0.06 0.14	0.08 0.19		0.04 0.17	0.06 0.06			0.10 0.21		0.06 0.16			0.06 0.15	0.06 0.17			0.07 0.07	
AA1*	0.08 0.19	0.06 0.14	0.04 0.17			0.07 0.07	0.10 0.21		0.06 0.16				0.07 0.18	0.06 0.16				0.06 0.06
BA1					0.05 0.15						0.05 0.16						0.05 0.11	
AA1						0.06 0.18						0.05 0.17						0.05 0.11
BC1					0.04 0.12	0.05 0.14					0.05 0.17	0.03 0.13					0.06 0.14	0.05 0.12
AC1					0.04 0.14						0.04 0.12	0.05 0.17					0.05 0.12	0.05 0.14

For each Hawkes effect, this figure presents the daily average branching ratio and adjusted branching ratio.

Panel II: Trades and LOS w/ BPI effect on LOB events

	BMW											SAP											ADS										
	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11
Buy*	0.45	0.34	0.31	0.22	0.18							0.44	0.36	0.33	0.16	0.11							0.40	0.35	0.38	0.24	0.18						
	0.06	0.06	0.04	0.03	0.05							0.06	0.05	0.03	0.02	0.03							0.06	0.07	0.06	0.04	0.05						
Sell*	0.32	0.14	0.26	0.17	0.23	0.31		0.33	0.41	0.27	0.08	0.39	0.06	0.36	0.30	0.21	0.16	0.20	0.20	0.13	0.42	0.10	0.25	0.09	0.25	0.14	0.22			0.18	0.28	0.30	0.10
	0.04	0.02	0.03	0.02	0.06	0.14		0.24	0.36	0.13	0.14	0.05	0.01	0.04	0.04	0.05	0.08	0.14	0.25	0.16	0.25	0.20	0.04	0.02	0.04	0.02	0.06			0.11	0.21	0.13	0.16
Buy	0.19	0.13	0.15									0.13	0.09	0.11									0.20	0.13	0.11								
	0.01	0.01	0.01									0.01	0.01	0.01									0.02	0.02	0.01								
Sell											0.07	0.08										0.06											0.08
											0.06	0.01										0.12											0.07
BA1*	0.17	0.16	0.29	0.24	0.17	0.10						0.12	0.17	0.48	0.29	0.18	0.10	0.08					0.15	0.11	0.28	0.23	0.12	0.14	0.10				
	0.05	0.06	0.08	0.08	0.11	0.11						0.03	0.05	0.10	0.08	0.09	0.10	0.10					0.06	0.06	0.11	0.10	0.09	0.13	0.12				
AA1*			0.04																	0.03													
			0.01																	0.08													
	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11
Buy*	0.34	0.06	0.28	0.18	0.21	0.30		0.34	0.40	0.27	0.08	0.42	0.14	0.33	0.26	0.17	0.21	0.26	0.25	0.13	0.42	0.09	0.24	0.08	0.26	0.15	0.23		0.28	0.17	0.29	0.30	0.09
	0.04	0.01	0.03	0.03	0.06	0.14		0.25	0.37	0.13	0.15	0.06	0.02	0.03	0.03	0.04	0.10	0.18	0.28	0.15	0.26	0.21	0.04	0.02	0.04	0.03	0.06		0.13	0.11	0.22	0.13	0.16
Sell*	0.45	0.34	0.32	0.20	0.19							0.45	0.37	0.34	0.17	0.13	0.08						0.44	0.37	0.41	0.25	0.20	0.19					
	0.06	0.05	0.04	0.03	0.05							0.06	0.05	0.03	0.02	0.03	0.03						0.07	0.07	0.06	0.04	0.05	0.07					
Buy												0.13																					
												0.01																					
Sell	0.20	0.12	0.10									0.15	0.10										0.28	0.13	0.18								
	0.01	0.01	0.01									0.01	0.01										0.02	0.01	0.01								
BA1*			0.05										0.05							0.04													
			0.01										0.02							0.09													
AA1*	0.17	0.17	0.32	0.23	0.16	0.09						0.14	0.19	0.49	0.28	0.20	0.11						0.15	0.12	0.27	0.22	0.11	0.13	0.09				
	0.05	0.06	0.09	0.08	0.11	0.10						0.04	0.06	0.10	0.08	0.08	0.10						0.06	0.06	0.11	0.10	0.08	0.12	0.10				
	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11
Buy*		0.12	0.13	0.15	0.14	0.14			0.13	0.15	0.19			0.10	0.16	0.17	0.14	0.14			0.14	0.12		0.16	0.17	0.15	0.16	0.11			0.13	0.16	0.20
		0.01	0.02	0.03	0.05	0.05			0.13	0.15	0.09			0.02	0.03	0.04	0.05	0.10			0.21	0.07		0.02	0.03	0.03	0.05	0.04			0.11	0.14	0.09
Sell*	0.80	0.84	0.55	0.24	0.19	0.11	0.14	0.11	0.08			0.87	0.94	0.43	0.22	0.18	0.16	0.10	0.10				0.75	0.82	0.49	0.27	0.23	0.16	0.09	0.12	0.08		
	0.11	0.10	0.09	0.05	0.06	0.04	0.07	0.09	0.08			0.11	0.10	0.07	0.04	0.05	0.06	0.07	0.10				0.11	0.12	0.09	0.06	0.07	0.05	0.05	0.08	0.07		
Buy																																	
Sell	0.20	0.16	0.16	0.13	0.09	0.08	0.08	0.07				0.16	0.18	0.13	0.08	0.07							0.21	0.15	0.15	0.10		0.09					
	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.03				0.01	0.01	0.01	0.01	0.01							0.02	0.01	0.01	0.01		0.02					
BA1*		0.21	0.10	0.25	0.12	0.36	0.28	0.17	0.15	0.17	0.61		0.23	0.10	0.30	0.19	0.34	0.23	0.16	0.10	0.09	0.59		0.17	0.10	0.22	0.10	0.22	0.19	0.15	0.12	0.15	0.61
		0.06	0.04	0.12	0.09	0.32	0.33	0.32	0.33	0.41	0.65		0.05	0.03	0.11	0.10	0.26	0.33	0.35	0.37	0.29	0.71		0.07	0.05	0.13	0.08	0.21	0.25	0.27	0.27	0.35	0.66
AA1*	0.29	0.17	0.15	0.10	0.07	0.04						0.31	0.23	0.14	0.07	0.07							0.27	0.18	0.17	0.10	0.08	0.06	0.04				
	0.09	0.05	0.05	0.05	0.05	0.04						0.08	0.05	0.04	0.03	0.04							0.10	0.07	0.08	0.06	0.06	0.06	0.06				
	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11
Buy*	0.81	0.85	0.52	0.22	0.18	0.10	0.13	0.11	0.08			0.83	0.92	0.41	0.23	0.16	0.15	0.11					0.74	0.79	0.46	0.25	0.20	0.15	0.09	0.12	0.08		
	0.11	0.10	0.08	0.05	0.06	0.04	0.07	0.08	0.07			0.11	0.10	0.06	0.04	0.04	0.06	0.07					0.11	0.12	0.08	0.06	0.06	0.05	0.05	0.08	0.07		
Sell*		0.14	0.12	0.14	0.15	0.14			0.11	0.16	0.19		0.15	0.13	0.16	0.17	0.15				0.15	0.13		0.16	0.18	0.20	0.17	0.11			0.14	0.16	0.21
		0.02	0.02	0.03	0.05	0.05			0.11	0.16	0.09		0.01	0.02	0.03	0.04	0.05				0.22	0.08		0.02	0.03	0.05	0.05	0.04			0.12	0.14	0.09
Buy	0.20	0.15	0.14	0.11	0.08	0.07						0.17	0.16	0.11	0.12	0.07	0.07						0.21	0.13	0.14	0.09		0.11					
	0.01	0.01	0.01	0.01	0.01	0.01						0.02	0.01	0.01	0.01	0.01	0.02						0.02	0.01	0.01	0.01		0.02					
Sell																																	
BA1*	0.30	0.18	0.15	0.09	0.06	0.04						0.29	0.21	0.12	0.07	0.09							0.25	0.19	0.17	0.10	0.07	0.07	0.05				
	0.09	0.05	0.06	0.04	0.05	0.04						0.08	0.05	0.04	0.02	0.04							0.10	0.07	0.08	0.06	0.06	0.06	0.07				
AA1*		0.21	0.10	0.25	0.11	0.35	0.28	0.18	0.13	0.16	0.60		0.24	0.11	0.27	0.21	0.37	0.29	0.17	0.11	0.09	0.60		0.16	0.10	0.24	0.10	0.21	0.19	0.16	0.12	0.14	0.61
		0.06	0.04	0.12	0.09	0.32	0.34	0.33	0.29	0.39	0.65		0.05	0.03	0.09	0.10	0.27	0.37	0.35	0.37	0.28	0.71		0.06	0.05	0.15	0.09	0.20	0.25	0.28	0.27	0.35	0.66

Panel III: LOB Events effects on Bid LOS w/o BPI

	BMW										
	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11
BA1	0.08 0.08	0.06 0.07	0.07 0.06	0.04 0.04							
BA2	0.13 0.11	0.07 0.07	0.08 0.06								
BA3		0.06 0.08	0.08 0.08	0.10 0.12							
BA4					0.08 0.15						
BA5				0.10 0.05		0.09 0.15					
BA6							0.18 0.23				
BA7								0.14 0.18			
BA8									0.19 0.24		
BA9								0.14 0.11		0.36 0.19	
AA1											
AA2											
AA3			0.04 0.04								
AA4											
AA5											
AA6											
AA7											
AA8											
AA9											

	SAP										
	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11
BA1	0.09 0.09	0.07 0.07	0.05 0.04								
BA2	0.12 0.12		0.06 0.04								
BA3		0.05 0.06	0.07 0.07	0.05 0.07							
BA4					0.05 0.08						
BA5											
BA6							0.09 0.13				
BA7								0.12 0.20			
BA8									0.15 0.15		
BA9										0.33 0.18	
AA1											
AA2											
AA3											
AA4											
AA5											
AA6											
AA7											
AA8											
AA9											

	ADS										
	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11
BA1	0.07 0.07	0.05 0.06	0.06 0.06	0.03 0.03							
BA2		0.04 0.04	0.07 0.05								
BA3		0.05 0.06	0.05 0.05	0.07 0.08							
BA4					0.07 0.11						
BA5						0.09 0.11					
BA6					0.07 0.05		0.11 0.14				
BA7								0.11 0.15			
BA8									0.17 0.20		
BA9										0.30 0.18	
AA1											
AA2											
AA3											
AA4											
AA5											
AA6											
AA7											
AA8											
AA9											

Panel III: LOB Events effects on Bid LOS w/o BPI (cont.)

	BMW										
	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11
BC1	0.10	0.03	0.06							0.05	
	0.10	0.04	0.06							0.17	
BC2	0.07	0.02	0.06	0.06							
	0.08	0.03	0.06	0.07							
BC3		0.04	0.06	0.11						0.05	
		0.04	0.04	0.10						0.15	
BC4				0.08	0.08						0.04
				0.05	0.10						0.33
BC5				0.08	0.23						
				0.04	0.19						
BC6				0.22		0.20					
				0.09		0.24					
BC7							0.15				
							0.16				
BC8								0.11			
								0.10			
BC9											
AC1	0.09	0.05	0.07	0.05							
	0.09	0.07	0.07	0.06							
AC2	0.06	0.04	0.05								
	0.06	0.05	0.05								
AC3	0.09	0.06	0.09	0.05							
	0.08	0.06	0.07	0.04							
AC4											
AC5											
AC6											
AC7											
AC8											
AC9											
AC10											

	SAP										
	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11
BC1	0.10	0.03	0.07	0.04		0.03				0.04	
	0.11	0.03	0.05	0.04		0.11				0.20	
BC2	0.08	0.12	0.05	0.04					0.03		
	0.10	0.16	0.05	0.06					0.32		
BC3		0.06	0.31	0.11	0.05						
		0.06	0.21	0.09	0.08						
BC4				0.43	0.07	0.04					
				0.34	0.09	0.10					
BC5					0.40						
					0.37						
BC6				0.15		0.31					
				0.05		0.38					
BC7							0.12				
							0.13				
BC8								0.10			
								0.11			
BC9											
AC1	0.08	0.07	0.07	0.04							
	0.08	0.07	0.05	0.04							
AC2	0.07	0.07	0.06								
	0.09	0.09	0.06								
AC3		0.14	0.12								
		0.13	0.08								
AC4											
AC5											
AC6											
AC7											
AC8											
AC9											
AC10											

	ADS										
	BA1	BA2	BA3	BA4	BA5	BA6	BA7	BA8	BA9	BA10	BA11
BC1	0.11	0.04	0.07	0.03				0.03		0.06	
	0.11	0.06	0.07	0.03				0.14		0.19	
BC2	0.06	0.03	0.06	0.05				0.04	0.05		
	0.07	0.05	0.06	0.06				0.16	0.25		
BC3		0.03		0.14						0.05	
		0.04		0.12						0.13	
BC4					0.06						0.04
					0.07						0.29
BC5				0.08	0.21	0.07					
				0.04	0.18	0.08					
BC6				0.19		0.36					
				0.09		0.35					
BC7							0.25				
							0.23				
BC8								0.13			
								0.11			
BC9											
AC1	0.09	0.05	0.08	0.05	0.04						
	0.09	0.07	0.08	0.06	0.08						
AC2	0.08	0.04	0.05	0.03							
	0.08	0.06	0.05	0.04							
AC3	0.11	0.05	0.06	0.05							
	0.09	0.06	0.05	0.04							
AC4				0.05							
				0.03							
AC5											
AC6											
AC7											
AC8											
AC9											
AC10											

Panel IV: LOB Events effects on Ask LOS w/ BPI

BMW

	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11
BA1											
BA2											
BA3			0.04 0.04								
BA4											
BA5											
BA6											
BA7											
BA8											
BA9											
AA1	0.08 0.08	0.06 0.07	0.07 0.06	0.03 0.04							
AA2	0.10 0.08	0.07 0.07	0.08 0.06								
AA3	0.10 0.11	0.06 0.08	0.08 0.08	0.08 0.10							
AA4					0.07 0.13						
AA5				0.09 0.05		0.11 0.18					
AA6							0.16 0.20				
AA7								0.16 0.20			
AA8									0.18 0.22		
AA9										0.36 0.18	

SAP

	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11
BA1											
BA2											
BA3			0.03 0.03								
BA4											
BA5											
BA6											
BA7											
BA8											
BA9											
AA1	0.10 0.10	0.07 0.07	0.05 0.04								
AA2	0.12 0.11		0.07 0.05								
AA3		0.04 0.06	0.07 0.07	0.05 0.06							
AA4					0.05 0.08						
AA5											
AA6											
AA7											
AA8									0.18 0.17		
AA9										0.29 0.16	

ADS

	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11
BA1											
BA2											
BA3											
BA4											
BA5											
BA6											
BA7											
BA8											
BA9											
AA1	0.09 0.09	0.05 0.06	0.06 0.05	0.03 0.03							
AA2		0.05 0.05	0.07 0.05								
AA3		0.05 0.07	0.05 0.05	0.08 0.09							
AA4					0.06 0.11						
AA5						0.09 0.11					
AA6							0.10 0.13				
AA7								0.11 0.15			
AA8									0.17 0.19		
AA9										0.30 0.17	

Panel IV: LOB events effects on Ask LOS w/o BPI (cont.)

BMW	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11
BC1	0.09 0.09	0.06 0.07	0.07 0.07	0.04 0.05							
BC2	0.06 0.06	0.05 0.06	0.05 0.05	0.03 0.04							
BC3	0.09 0.07	0.06 0.07	0.09 0.07	0.03 0.03							
BC4				0.05 0.04							
BC5											
BC6											
BC7											
BC8											
BC9											
AC1	0.11 0.11	0.03 0.04	0.06 0.06							0.05 0.17	
AC2		0.03 0.04	0.06 0.06	0.05 0.07							
AC3		0.04 0.04	0.06 0.05	0.10 0.10						0.04 0.14	
AC4				0.06 0.04	0.08 0.10						0.04 0.36
AC5				0.08 0.03	0.24 0.20						
AC6				0.22 0.09		0.17 0.21					
AC7							0.13 0.15				
AC8								0.11 0.11			
AC9											
AC10											

SAP	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11
BC1	0.08 0.09	0.06 0.07	0.07 0.05	0.04 0.04							
BC2	0.07 0.10	0.07 0.10	0.08 0.08								
BC3		0.13 0.12	0.14 0.09								
BC4											
BC5											
BC6											
BC7											
BC8											
BC9											
AC1	0.09 0.10	0.03 0.03	0.06 0.04	0.04 0.04						0.04 0.20	
AC2	0.07 0.09	0.10 0.14	0.05 0.05	0.04 0.05					0.03 0.32		
AC3		0.04 0.04	0.28 0.19	0.09 0.08	0.05 0.08						
AC4				0.41 0.33	0.06 0.08	0.05 0.15					
AC5				0.08 0.05	0.37 0.34						
AC6				0.15 0.05	0.11 0.06	0.27 0.33					
AC7							0.21 0.22				
AC8								0.18 0.19			
AC9											
AC10											

ADS	AA1	AA2	AA3	AA4	AA5	AA6	AA7	AA8	AA9	AA10	AA11
BC1	0.09 0.10	0.05 0.07	0.08 0.08	0.05 0.06	0.04 0.08						
BC2		0.04 0.06	0.05 0.05	0.04 0.04							
BC3	0.14 0.12	0.05 0.06	0.06 0.05	0.04 0.04							
BC4				0.04 0.03							
BC5											
BC6											
BC7											
BC8											
BC9											
AC1	0.10 0.10	0.04 0.06	0.08 0.07					0.03 0.13		0.07 0.19	
AC2	0.06 0.07	0.03 0.04	0.06 0.06	0.05 0.05				0.03 0.15	0.05 0.25		
AC3		0.03 0.04	0.05 0.04	0.14 0.13						0.05 0.13	
AC4				0.09 0.06	0.06 0.07						0.04 0.31
AC5				0.07 0.04	0.19 0.16	0.07 0.08					
AC6				0.17 0.08	0.08 0.06	0.31 0.30					
AC7							0.22 0.21				
AC8								0.11 0.10			
AC9											
AC10											

Panel V: LOB events effects on Bid LOC

	BMW										
	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11
BA1	0.13 0.13	0.03 0.03		0.03 0.05							
BA2		0.06 0.04									
BA3			0.07 0.09								
BA4											
BA5					0.17 0.20						
BA6											
BA7											
BA8											
BA9											
AA1	0.06 0.06	0.05 0.05	0.05 0.06								
AA2	0.04 0.03	0.04 0.03	0.04 0.04								
AA3	0.07 0.07	0.06 0.06	0.06 0.07								
AA4											
AA5											
AA6											
AA7											
AA8											
AA9											

	SAP										
	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11
BA1	0.19 0.19	0.03 0.02	0.03 0.03								
BA2		0.17 0.13									
BA3			0.15 0.22								
BA4				0.23 0.28							
BA5					0.32 0.34						
BA6						0.34 0.27					
BA7											
BA8											
BA9											
AA1	0.07 0.06	0.06 0.04	0.03 0.04								
AA2	0.05 0.05	0.09 0.06									
AA3		0.09 0.09									
AA4											
AA5											
AA6											
AA7											
AA8											
AA9											

	ADS										
	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11
BA1	0.15 0.15	0.03 0.03									
BA2											
BA3											
BA4											
BA5											
BA6											
BA7											
BA8											
BA9											
AA1	0.06 0.06	0.05 0.04	0.04 0.05								
AA2	0.04 0.03	0.05 0.03	0.06 0.05								
AA3	0.05 0.05	0.06 0.06	0.06 0.08								
AA4											
AA5											
AA6											
AA7											
AA8											
AA9											

Panel V: LOB events effects on Bid LOC (cont.)

BMW	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11
BC1	0.17	0.20	0.10	0.05	0.03						
	0.17	0.18	0.12	0.07	0.09						
BC2	0.08	0.10	0.06	0.03							
	0.09	0.10	0.08	0.06							
BC3	0.07	0.09	0.08	0.06	0.04	0.03					
	0.06	0.07	0.08	0.08	0.08	0.08					
BC4				0.06	0.04	0.05					
				0.06	0.07	0.10					
BC5				0.14		0.07					
				0.09		0.09					
BC6					0.08		0.06				
					0.07		0.08				
BC7							0.13				
							0.13				
BC8											
BC9											
AC1		0.02		0.04							
		0.02		0.06							
AC2											
AC3			0.04								
			0.04								
AC4											
AC5											
AC6											
AC7											
AC8											
AC9											
AC10											

SAP	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11
BC1	0.15	0.18	0.06	0.04							
	0.15	0.14	0.08	0.05							
BC2	0.06	0.09	0.06	0.05							
	0.08	0.09	0.09	0.09							
BC3	0.10	0.07	0.09	0.05	0.07						
	0.08	0.05	0.09	0.06	0.12						
BC4											
BC5				0.07							
				0.04							
BC6											
BC7											
BC8											
BC9											
AC1				0.04							
				0.05							
AC2											
AC3			0.10								
			0.11								
AC4											
AC5											
AC6											
AC7											
AC8											
AC9											
AC10											

ADS	BC1	BC2	BC3	BC4	BC5	BC6	BC7	BC8	BC9	BC10	BC11
BC1	0.16	0.18	0.10	0.05	0.04						
	0.16	0.18	0.12	0.07	0.09						
BC2	0.07	0.08	0.07	0.04	0.03						
	0.07	0.08	0.09	0.06	0.07						
BC3	0.07	0.07	0.08	0.06	0.04	0.04					
	0.06	0.06	0.08	0.08	0.06	0.08					
BC4				0.07	0.07	0.05	0.04				
				0.07	0.09	0.07	0.08				
BC5				0.06		0.05	0.04				
				0.05		0.06	0.06				
BC6							0.05				
							0.06				
BC7											
BC8											
BC9											
AC1				0.03							
				0.05							
AC2											
AC3											
AC4											
AC5											
AC6											
AC7											
AC8											
AC9											
AC10											

Panel VI: LOB events effects on Ask LOC

BMW	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11
BA1	0.07 0.07	0.06 0.05	0.05 0.06	0.03 0.04							
BA2	0.05 0.04	0.04 0.03	0.04 0.03								
BA3		0.06 0.06	0.06 0.08								
BA4											
BA5											
BA6											
BA7											
BA8											
BA9											
AA1	0.13 0.13	0.03 0.03		0.03 0.05							
AA2		0.05 0.04									
AA3			0.07 0.09								
AA4											
AA5					0.17 0.20						
AA6											
AA7											
AA8											
AA9											

SAP	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11
BA1	0.07 0.07	0.06 0.05	0.04 0.04								
BA2	0.04 0.04	0.09 0.06									
BA3		0.08 0.08	0.10 0.15								
BA4											
BA5											
BA6											
BA7											
BA8											
BA9											
AA1	0.17 0.17										
AA2		0.17 0.12									
AA3			0.15 0.22								
AA4				0.23 0.28							
AA5					0.31 0.33						
AA6						0.34 0.27					
AA7											
AA8											
AA9											

ADS	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11
BA1	0.06 0.06	0.05 0.05	0.04 0.05								
BA2	0.04 0.03	0.05 0.03	0.06 0.05								
BA3	0.06 0.05	0.05 0.05	0.06 0.07								
BA4											
BA5											
BA6											
BA7											
BA8											
BA9											
AA1	0.16 0.15										
AA2											
AA3											
AA4											
AA5											
AA6											
AA7											
AA8											
AA9											

Panel VI: LOB events effects on Ask LOC (cont.)

BMW	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11
BC1		0.03 0.02		0.03 0.05							
BC2											
BC3			0.04 0.05								
BC4											
BC5											
BC6											
BC7											
BC8											
BC9											
AC1	0.18 0.18	0.20 0.18	0.10 0.12	0.04 0.06	0.03 0.09						
AC2	0.09 0.10	0.10 0.10	0.06 0.08	0.03 0.05							
AC3	0.09 0.07	0.09 0.07	0.09 0.09	0.05 0.07	0.04 0.08	0.03 0.08					
AC4				0.08 0.08	0.05 0.08	0.05 0.10					
AC5				0.15 0.09		0.07 0.09					
AC6					0.08 0.07		0.06 0.08				
AC7							0.14 0.14				
AC8											
AC9											
AC10									0.15 0.15		

SAP	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11
BC1				0.04 0.05							
BC2											
BC3											
BC4				0.09 0.09							
BC5											
BC6											
BC7											
BC8											
BC9											
AC1	0.15 0.15	0.19 0.14	0.07 0.08	0.05 0.06							
AC2	0.07 0.09	0.10 0.10	0.06 0.09	0.05 0.07							
AC3	0.10 0.09	0.07 0.05	0.09 0.09	0.04 0.05	0.07 0.11						
AC4					0.04 0.06						
AC5											
AC6											
AC7											
AC8											
AC9											
AC10											

ADS	AC1	AC2	AC3	AC4	AC5	AC6	AC7	AC8	AC9	AC10	AC11
BC1											
BC2											
BC3											
BC4											
BC5											
BC6											
BC7											
BC8											
BC9											
AC1	0.17 0.17	0.19 0.19	0.10 0.12	0.05 0.07	0.04 0.09						
AC2	0.07 0.07	0.08 0.08	0.07 0.09	0.03 0.05	0.03 0.07						
AC3	0.07 0.06	0.07 0.06	0.09 0.09	0.05 0.07	0.04 0.07	0.04 0.08					
AC4				0.07 0.07	0.07 0.09	0.04 0.07	0.04 0.08				
AC5				0.07 0.05		0.06 0.06	0.04 0.06				
AC6							0.04 0.05				
AC7											
AC8											
AC9											
AC10											

Table 4.6 Multi-events mutual-excitation examples

Panel I: Levels 1 tot 3 Bid LOS w/o BPI and Ask LOC

BMW						
	BA1	BA2	BA3	AC1	AC2	AC3
BA1	0.08	0.06	0.07	0.07	0.06	0.05
	0.08	0.07	0.06	0.07	0.05	0.06
BA2	0.13	0.07	0.08	0.05	0.04	0.04
	0.11	0.07	0.06	0.04	0.03	0.03
BA3		0.06	0.08		0.06	0.06
		0.08	0.08		0.06	0.08
AC1	0.09	0.05	0.07	0.18	0.20	0.10
	0.09	0.07	0.07	0.18	0.18	0.12
AC2	0.06	0.04	0.05	0.09	0.10	0.06
	0.06	0.05	0.05	0.10	0.10	0.08
AC3	0.09	0.06	0.09	0.09	0.09	0.09
	0.08	0.06	0.07	0.07	0.07	0.09

SAP						
	BA1	BA2	BA3	AC1	AC2	AC3
BA1	0.09	0.07	0.05	0.07	0.06	0.04
	0.09	0.07	0.04	0.07	0.05	0.04
BA2	0.12		0.06	0.04	0.09	
	0.12		0.04	0.04	0.06	
BA3		0.05	0.07		0.08	0.10
		0.06	0.07		0.08	0.15
AC1	0.08	0.07	0.07	0.15	0.19	0.07
	0.08	0.07	0.05	0.15	0.14	0.08
AC2	0.07	0.07	0.06	0.07	0.10	0.06
	0.09	0.09	0.06	0.09	0.10	0.09
AC3		0.14	0.12	0.10	0.07	0.09
		0.13	0.08	0.09	0.05	0.09

ADS						
	BA1	BA2	BA3	AC1	AC2	AC3
BA1	0.07	0.05	0.06	0.06	0.05	0.04
	0.07	0.06	0.06	0.06	0.05	0.05
BA2		0.04	0.07	0.04	0.05	0.06
		0.04	0.05	0.03	0.03	0.05
BA3		0.05	0.05	0.06	0.05	0.06
		0.06	0.05	0.05	0.05	0.07
AC1	0.09	0.05	0.08	0.17	0.19	0.10
	0.09	0.07	0.08	0.17	0.19	0.12
AC2	0.08	0.04	0.05	0.07	0.08	0.07
	0.08	0.06	0.05	0.07	0.08	0.09
AC3	0.11	0.05	0.06	0.07	0.07	0.09
	0.09	0.06	0.05	0.06	0.06	0.09

Panel II: Levels 1 tot 3 Ask LOS w/o BPI and Bid LOC

BMW						
	AA1	AA2	AA3	BC1	BC2	BC3
AA1	0.08	0.06	0.07	0.06	0.05	0.05
	0.08	0.07	0.06	0.06	0.05	0.06
AA2	0.10	0.07	0.08	0.04	0.04	0.04
	0.08	0.07	0.06	0.03	0.03	0.04
AA3	0.10	0.06	0.08	0.07	0.06	0.06
	0.11	0.08	0.08	0.07	0.06	0.07
BC1	0.09	0.06	0.07	0.17	0.20	0.10
	0.09	0.07	0.07	0.17	0.18	0.12
BC2	0.06	0.05	0.05	0.08	0.10	0.06
	0.06	0.06	0.05	0.09	0.10	0.08
BC3	0.09	0.06	0.09	0.07	0.09	0.08
	0.07	0.07	0.07	0.06	0.07	0.08

SAP						
	AA1	AA2	AA3	BC1	BC2	BC3
AA1	0.10	0.07	0.05	0.07	0.06	0.03
	0.10	0.07	0.04	0.06	0.04	0.04
AA2	0.12		0.07	0.05	0.09	
	0.11		0.05	0.05	0.06	
AA3		0.04	0.07		0.09	
		0.06	0.07		0.09	
BC1	0.08	0.06	0.07	0.15	0.18	0.06
	0.09	0.07	0.05	0.15	0.14	0.08
BC2	0.07	0.07	0.08	0.06	0.09	0.06
	0.10	0.10	0.08	0.08	0.09	0.09
BC3		0.13	0.14	0.10	0.07	0.09
		0.12	0.09	0.08	0.05	0.09

ADS						
	AA1	AA2	AA3	BC1	BC2	BC3
AA1	0.09	0.05	0.06	0.06	0.05	0.04
	0.09	0.06	0.05	0.06	0.04	0.05
AA2		0.05	0.07	0.04	0.05	0.06
		0.05	0.05	0.03	0.03	0.05
AA3		0.05	0.05	0.05	0.06	0.06
		0.07	0.05	0.05	0.06	0.08
BC1	0.09	0.05	0.08	0.16	0.18	0.10
	0.10	0.07	0.08	0.16	0.18	0.12
BC2		0.04	0.05	0.07	0.08	0.07
		0.06	0.05	0.07	0.08	0.09
BC3	0.14	0.05	0.06	0.07	0.07	0.08
	0.12	0.06	0.05	0.06	0.06	0.08

Chapter 5
On limit orders and quote-stuffing activities

In the previous chapter, we have performed an extensive analysis of the limit order book (LOB) events and their interrelations. This comprehensive work was made possible by the data obtained using the events identification methodology developed in Chapter 2. However, we have to keep in mind that none of these events would be observed if it was not for the limit orders, which come and go across the book according to the will of the market participants. Therefore, in this fifth and final chapter, we will focus on the limit orders themselves, more particularly on those whose lifecycle ends with a cancellation. Indeed, while much work has focused on the actual trades, and therefore, on the aggressive orders and the limit orders eventually executed, the canceled limit orders have received less attention. We will begin by considering these liquidity vehicles individually and then, we will look at those appearing submitted in sequences. In order to perform our analysis, we will use the Xetra data processed using the methodologies developed in Chapters 1 to 3. In fact, because we mainly focus on limit orders, we are able to take advantage of the data produced using the orders tracking methodology developed in Chapter 3. Section 5.1 presents detailed information regarding this data and its context.

Once the data presented, this chapter is divided into three main parts. First, in section 5.2, we work on the limit orders by considering them individually. Through multiple general observations, the main objective of this section is to relate the behaviors observed through limit orders submission and cancellation to the traditional definition of an order-driven order book in which patient investors submit passive orders with an actual execution goal in mind. Therefore, in sub-section 5.2.1, we relate the orders submission and cancellation contexts by analyzing their movements in the book during their life. In sub-section 5.2.2, we work on the limit orders duration, which we define as the time elapsed between their submission and their cancellation. Finally, in sub-section 5.2.3, we look at the limit orders execution actual potential.

In the second part of this chapter, section 5.3 focus on quote-stuffing activities. The works of Egginton, Van Ness et al. (2016) represent the most important reference on the matter.

These authors define quote stuffing as “the practice where a large number of orders to buy or sell securities are placed and then canceled almost immediately”. The main goal of these activities is to overload the stock market as well as the other traders systems in order to create latency arbitrage opportunities. It appears clear that the arrival of High-Frequency Trading (HFT) algorithms have been directly related to the rise of such practices. While new regulations have tried to discourage these activities in more recent years, it was not the case during our 2013 Xetra dataset period. This lack of regulation and the important technological advances observed in the preceding years lead us to believe that our data comes from a period that was particularly favorable for quote-stuffing activities.

It is important to mention that our objective is not to question the consistency of the Egginton, Van Ness et al. (2016) methodology as well as the quality of their conclusions, but to examine quote-stuffing activities from another point of view. These authors have mostly analyzed the phenomenon with inter-markets, inter-stocks, and pre/post quote-stuffing market quality concerns, which we will leave aside to work on the sequences of submitted and cancelled limit orders that we relate to quote-stuffing operations. Because of the detailed limit orders information that we have been able to obtain through our incremental data processing methodologies developed in chapters 1 to 3, we are in a very good position to observe these activities at a microscopic level. Therefore, the main contribution of this chapter lies in the identification and characterization of different algorithms performing quote-stuffing operations. In section 5.3.1, we first implement a methodology allowing to identify the limit orders sequences potentially resulting from these activities. Indeed, because we still have no information regarding individual orders owner and the fact that they are submitted through an algorithm or not, once again we have to go over some deductions in order to determine their potential to be associated with quote-stuffing. In section 5.3.2, by focusing on their characteristics as well as their submission and cancellation patterns, we analyze the behavior of the algorithms that potentially have submitted and cancelled these orders. We develop a frame of reference allowing us to classify the observed results based on different characteristics such as the number of physical prices noticed in the involved orders.

5.1 The data

In Chapter 3, we have developed a limit orders tracking methodology and used it on the DAX, MDAX and SDAX stock indexes components for the 61 trading days time period spanning from February 1st to April 30th, 2013. These three indexes encompass the stocks of 30 large, 50 medium and 50 small market capitalizations from non-technologic sectors. We have mainly focused our algorithmic performances when applied to these stocks. In the current chapter, we intend to use this data in the context of economic analyzes through which the behavior of some market participants will be analyzed. Most of our work will therefore focus on orders that have been qualified as successfully tracked in Chapter 3. However, before focussing on these orders, we consider relevant to present an overview of the Xetra market activity during the period of interest. In order to obtain this big picture, we use the data produced using Chapter 2 events identification methodology. Indeed, given the low loss rate resulting from this methodology regarding the limit orders submission and cancellation, this data appears quite suitable for this first task.

Figure 5.1 presents an aggregation of the initial price level for the 198.4, 72.0 and 31.7 millions of passive orders whose submission have been observed for the DAX, MDAX and SDAX components during our three-month period. We define the price level of an order as its position relative to its book side best visible price. The price level 1 corresponds to the best available price, the price level 2 to the second best price and so on. Figure 5.1 shows that orders submissions are highly concentrated on the five best bid or ask prices. Indeed, 74%, 71% and 81% of these activities appear to take place on these price levels. It is interesting to observe that while more distributed for the DAX and MDAX stocks, 68% of the small-cap SDAX stocks submissions are performed on the best and second best ask and bid price levels with an important peak of 57% for the sole first level. On the other hand, we note that with 7%, 13% and 4%, the proportions of orders submissions taking place on price levels 11 to 20 may be qualified as marginal.

Figure 5.2 presents similar information regarding the 187.4, 68.2 and 31.2 millions of passive orders cancellations observed for the DAX, MDAX and SDAX components. With 69%, 67% and 81%, these activities also appear concentrated on the first five price levels. However, from the 2nd to the 8th price level, the DAX and MDAX order cancellations proportions show a smoother descending trend compared to that observed for the submissions. Regarding the SDAX component, it is interesting to observe that most of the activity now takes place on the three best price levels with a new marked peak on the second one. Despite higher for the DAX and MDAX stocks, 12%, 14% and 3% proportions of orders cancellations occurring on depth levels 11 to 20 also appear marginal.

We find interesting to observe that compared to the orders submissions, the cancellations proportions exhibit a general shift toward the higher price levels. We relate this phenomenon to the potential patient investors behavior who would wait for their own orders to move away from the best prices before triggering their cancellation. We also consider important to note the 11.0, 3.7 and 0.5 million deviations between the global numbers of observed orders submissions and cancellations. We relate these differences to the executed orders and those for which we do not observe the cancellation. This second category consists in the orders remaining on the order book at least until the end of the trading day and those whose cancellation takes place outside the 20 price levels window of order book data actually provided by Xetra.

With this global picture in mind, it becomes more natural to shift our attention toward the orders that have been successfully tracked using our Chapter 3 methodology. At this point, we find important to recall that on average, these orders represent 41.5%, 87.4% and 98.6% of all orders whose submission have been observed for the DAX, MDAX and SDAX components. When it comes to the DAX index components, this smaller rate leads us to fear a selection bias that could emerge from the sole analysis of the successfully tracked orders. In order to mitigate this potential bias, we choose to work with the 15 stocks having presented the highest tracking success rates, leaving the others aside. Consequently, from this point, we refer to this stock group as *DAX15*. These stocks successfully tracked orders proportions span from 44.7% to 78.5%, with an average of 54.9%. We consider this selection relevant since as seen in Chapter 3 Figure 3.2 Panel I, these stocks appear equally distributed with regard to the general order book activity. Although we do not have information indicating that they are not representative of all orders, we consider that the statistics obtained from these stocks successfully tracked orders should be interpreted with caution. This concern should decrease in importance regarding the MDAX and SDAX indexes for which we keep working with all the constituents. Indeed, because of their more important success rates, we believe that the information obtained using their successfully tracked orders should be generally in line with those of the whole population. This assumption particularly applies to the SDAX index components and their 1.42% average rate of lost orders.

Disregarding the orders whose track is lost because they leave the 20 price levels window, we are in the presence of 57.1, 60.8 and 31.1 million of successfully tracked passive orders for the *DAX15*, MDAX and SDAX components. Globally, 6.8%, 4.3% and 1.3% of these orders are ultimately totally executed, which leads to cancellation rates of 93.2%, 95.7% and 98.7%. Even adding the fact that we observe the partial execution of 0.7%, 0.5% and 0.3% of these orders at some point before their cancellation, it is easy to conclude that most submitted limit orders end with a cancellation.

In order to analyze these outcomes according to their price level of submission, Table 5.1 presents the termination context of our successfully tracked passive orders. In addition to the proportions of totally executed orders and those whose track is lost because of the 20 price levels window, it relates the cancelled orders to their level of cancellation. First, we observe rapidly decreasing execution rates as the price level of submission increase. With the most important 19.3%, 11.1% and 2.0% rates of orders submitted on the best price level ending in a total execution, these values decrease to 0.7%, 0.6% and 0.2% for the orders entering the LOB on the 5th best price level. Second, we note that the proportion of orders lost because they end up outside the 20 depth levels window is increasing in price level of submission, reaching as far as values above 80% for the orders submitted on the 20th price level. We relate this observation to an increasing probability of reaching this limit for the orders submitted on price levels that are close to it.

5.2 On the limit orders

In this section, we focus on the life of the limit orders eventually ending with a cancellation. We first examine the movement of these orders through the order book between the time of their submission and the time of their cancellation. Second, we focus on the duration of the orders, which we define as the length of the time interval between the submission and the cancellation. Finally, on the basis of their duration, we analyze the actual potential for executing these orders.

5.2.1 Price movements before cancellation

Having presented the successfully tracked orders price level of submission and cancellation in section 5.1, it becomes interesting to link this information through the presentation the price levels of cancellation for each price level of submission. Table 5.1 reports this information for the *DAX15*, MDAX and SDAX components. This table shows important proportions of passive orders being cancelled on their actual price level of submission or immediate neighbors. This phenomenon appears particularly pronounced regarding the MDAX and SDAX. In the *DAX15* stocks case, these proportions appear to spread across a more important number of cancellation price level mainly above the submission price levels, especially regarding the order submitted on the second half of the LOB. This is consistent with the potential behavior of patient investors cancelling their limit orders presenting decreasing probabilities of execution. However, it is relevant to ask how much this information would change if we had access to these stocks lost orders. MDAX and SDAX stocks rates may not be as consistent with the patient investor expected behavior. Indeed, disregarding the orders submitted on the 20th price level, the proportions of passive order cancelled on their submission level range between 27.9% and 46.0% for the MDAX stocks. They reach even higher levels for the SDAX component, ranging between 32.2% and 69.1%. These observations are made interesting by the fact that these values remain high even ignoring those associated with the orders submitted on the two best price levels. Indeed, we do not necessarily expect the market participants who submit limit orders on these price levels to be patient investors. It is then not particularly surprising to observe that a very small proportion of these orders appears canceled above level 4.

Through the implementation of our Chapter 3 orders tracking methodology, we have been able to collect the number of position change with regard to the best price level over the life of each order. This information becomes highly relevant in the investigation of the important proportions of limit orders cancelled on their exact submission price level.

Thus, on a price level of submission basis, Figure 5.3 provides the proportion of share having been cancelled after zero, one, or more than one position change. The 20th price level is omitted because of its too close proximity to the limit visibility window made available by Xetra. Regarding the orders cancelled after a single price move, we make the distinction between a move toward the best price level (down move) and a move in the opposite direction (up move). It is important to note that unlike the proportions presented in the previous Table 5.1, the *zero move* and the *one move* categories exclude the orders that would have transited by different price levels before being cancelled on their price level of submission, one level under, or one level above. Although the proportions related to these orders vary among our three stock sets, they are generally significant, especially regarding the SDAX index components for which they represent the majority for all price levels below 16. Disregarding their arrival price level, these orders represent 47.6% and 54.3% of all orders ending with a cancellation or total execution for the *DAX15* and MDAX stocks. This proportion increases to 78.4% when it comes to the SDAX index components.

We question the intentions of the market participants who engage in this type of activity, especially when they take place on price levels 3 to 19. Although less important than for the two best price levels, for this deeper book section, its represents 36.6%, 40.2% and 71.1% of the submitted orders for which we ultimately observe the cancellation. We have to place a caveat on the orders cancelled after a single move away from the best price level since they could be related to some investors cancelling their orders after observing a first adverse movement in the price. On the other hand, we consider the cancellation of an order whose position with regard to the best price have not changed as inconsistent with a real execution intention. Our concern is even more important when it comes to a passive order having gained one position toward the best price level in a single move. Although not impossible for these cancellations to be related to events taking place on the other side of the order book, their large proportions force us to push the investigation further. Indeed, it is even possible to question the fact that these orders are actually monitored by their owners.

When the orders are cancelled after two position changes or more, for more representativeness, especially on the deepest levels, we have included the orders whose track have been lost because they have reached the limit of 20 price levels. Because of the fact that at least two position changes have mechanically been required by these orders to reach this level, we consider their addition to the set as legitimate for this part of the analysis. Thus, regarding all the orders present in this group, we consider the distribution of their proportions among the price levels of arrival as more consistent with the notion of investor patience. We claim that these orders involve a potentially higher level of patience than those previously described. In this context, their increasing proportions between the first and 10^{th} price levels of submission, may be related to the investors patience level that could be increasing with their favorite level of submission. Beyond this point, we observe a general stability, a slight increasing trend and a still decreasing trend for the *DAX15*, MDAX and SDAX components. The very low proportion of the total book activity taking part on these levels makes it difficult to explain these three different tendencies. We must exercise general caution in analyzing the Figure 5.3 proportions since the activity observed on lower price levels may lead to more position changes for the orders standing deeper in the book, which may affect the time before an order first move. Indeed, we have to recall that a price level $l \in \{1, \ldots, 19\}$ creation or deletion mechanically causes a position change for all the orders already standing on level $l + 1$ and above. In this context, we consider highly relevant to continue our investigation, especially, with the analysis of orders duration.

5.2.2 Limit orders duration

As explained before, we define a limit order duration as the time elapsed between its submission and it cancellation. Figure 5.4 presents the cumulative distribution functions (CDF) of the *DAX15*, MDAX and SDAX components cancelled orders duration. It first shows that 31%, 35% and 51% of the cancelled limit orders are cancelled after having been part of the book for less than one second. Then, it shows that 60%, 50% and 42% of orders durations lies between one second and one minute. This leads to the third fact that about 9%, 15% and 7% of all the orders that we follow between their submission and cancellation are cancelled after having lied on the order book for more than one minute.

We consider these three boundaries as relevant orders category delimiters to continue our analysis. Therefore, we first define the *short duration* orders as those having a duration that is smaller than one second, since Moallemi and Sağlam (2013) have established the human reaction time to about 500 ms. Figure 5.4 actually shows that 25%, 30% and 47% of the orders cancellations take place within this time period. It is then possible to relate these orders to purely algorithmic, and systematic activities, generally referred as High-frequency Trading (HFT). However, we find reasonable to extend this assumption to all the orders cancelled within 1 second. Indeed, although theoretically possible for these orders to belong to human market participants, we have no doubts on the systematic nature of their cancellation. Indeed, we consider highly possible for the cancellation of the short duration orders to be pre-programmed, especially given that more than 50% of them are actually cancelled after less than 100 ms. This implies that these orders are not subject to any monitoring by their owners. While it is not impossible that they are part of global strategies involving some form of patience, we attribute absolutely no patience to the individual orders standing in the short duration category. We also qualify the liquidity provided through this type of order as virtual, since it is almost unreachable by conventional traders. Second, we define the *medium duration* orders as those presenting a time between submission and cancellation lying between 1 second and 1 minute. We relate these orders to short term, potentially algorithmic trading strategies that may involve monitoring. We find difficult to establish the patience level related to an order cancelled between 1 second and 1 minute after its submission. However, unlike short duration orders, we consider that some level of patience may be involved, especially given that in an algorithmic world, price movements can occur very quickly. Although potentially challenging, the liquidity provided through these orders may be accessible to any investor. Finally, we qualify the orders whose duration is larger than 1 minute as *long duration* orders. We relate these orders to longer term strategies that may be elaborated by patient, possibly more conventional traders. We consider the liquidity provided through these orders as real and reachable by anyone. From our point of view, they actually represent the real sustainable source of liquidity for the conventional investor.

At this point, we consider important to relate the previously presented positions change with regard to the best price level with actual orders duration. Figure 5.5 gives a summary of the situation by providing the orders durations related to our previous position changes categories. As we could imagine, it is possible to observe that 51%, 53% and 59% of orders for which we have observed no or only one move are now classified in the short duration category. We also denote that 48%, 33% and 25% of these orders present a medium duration. We consider possible to associate the cancellation of these orders with a change in strategy by investors observing an absence of movement in the price execution possibility. Finally, also not surprising, Figure 5.5 shows that 89%, 88% and 77% of the orders for which we have observed two position changes or more end up in the medium and long duration orders categories.

As before, we find interesting to analyze orders duration in the context of their submission price level. Figure 5.6 visually provides this information for the *DAX15*, MDAX and SDAX components. Since they may affect some conclusions, particularly regarding the second half of the order book, we include the orders whose track have been lost after they have crossed the 20 price levels boundary. According to the available information, we distribute these orders among three categories. Those whose track have been lost after more than 1 minute have been added to the regular long duration orders category. On the other hand, we have assigned their own category to the orders whose track have been lost between one second and one minute after submission. The same applies to those for which this event has occurred after less than one second. For our three stocks group, Figure 5.6 shows a decreasing proportion of short duration orders submissions taking place in the first half of the order book. This trend is generally constant between levels 1 to 9, 1 to 7 and 1 to 4 for the *DAX15*, MDAX and SDAX stocks. The proportions related to short duration orders are reassigned to the medium and long duration orders. We consider that this phenomenon is consistent with an increasing patience level of the investors as we increase in the depth level of orders submission. The extent of this trend appears to be linked to the general level of activity observed on these price levels which, as shown before, is increasing as we move from the *DAX15* to the SDAX. Although relatively stable on the intermediate price levels for the MDAX and SDAX components, considering possible for a part of the b*etween 1 sec and 1 min – above 20* to end up in this category, the proportion of long duration orders appears increasing even in the deepest segment of the book, which is still consistent with an investor increasing patience level. The medium duration orders represent the actual majority on 16, 13 and 10 of the presented price levels of submission. We conclude that the short-term trading activities involving some patience almost prevails throughout the entire order book. On the other hand, while particularly important on the best price levels, the proportion of short duration orders submissions reach a floor value in the mid segment of the book. Afterward, they exhibit a slightly increasing trend up to the deepest levels for the DAX15 and MDAX components. The SDAX index stocks orders belonging to this category represent a generally higher and more constant proportion than those of the DAX15 and MDAX, and this, from the 4^{th} price level of submission. It is interesting to observe how short duration orders submission, which we associate to impatient investors strategies, are widespread even across the deeper sections of the order book, especially regarding the SDAX stock components. With this phenomenon in mind, in the next sub-section, we will continue our investigation with the analysis of the execution potential for the orders belonging in each duration category.

5.2.3 Limit orders execution potential

Given the previous observations, we consider relevant to ask if the execution is a real possible outcome, particularly regarding the short duration orders which sometime stand on the LOB for a very short time. To analyze this question, Figure 5.7 presents the executed orders duration which, we define as the time between an order submission and its total execution. In very general terms, it shows a very similar distribution shape regarding the orders executed within 100 milliseconds for our three groups of stocks. Afterward, we note a shift in the cumulative distribution functions for the DAX15, MDAX and SDAX components. We conclude in a global decreasing orders execution speed as we move from the most to the less active stocks.

Figure 5.7 shows that over our three month period, 57.4%, 48.3% and 35.5% of the orders have been executed after having lied on the order book for a time period going from 1 second to 1 minute. It also shows that 3.1%, 21.6% and 37.8% of the executed orders have done the same for a period larger than 1 minute. Given these numbers, it is possible to establish that the previously identifier long and medium duration orders exhibit a real execution possibility. It is certain that this one may vary according to the price level of submission and the duration of each order, but the existence of this possibility is beyond doubt. On the other hand, this figure first shows that regarding the orders submitted on the best price level ultimately ending in a total execution, 14.1%, 14.0% and 13.1% are executed within a 10 milliseconds time span. These proportions becomes 9.4%, 8.2%, 7.7% for the orders executed after a time period going from 10 milliseconds and 100 milliseconds. When it comes to the orders executed after a 100 millisecond to 1 second period, they represent 15.9%, 7.9% and 5.7% of all the orders that we track between the submission and the total execution. In these specific cases, although relatively important, these rates cannot be related to our short duration orders execution possibility without taking the orders submission price level into consideration. In this context, through Figure 5.8, we redistribute this section of the cumulative distribution functions with regard to the orders price level of arrival. This figure first shows that 42.9%, 34.0% and 28.8% of the total executions of orders submitted on the best price level take place less than 1 second after the order submission. These proportions are composed of 16.4%, 16.3% and 14.9% of execution occurring in less than 10 milliseconds, 10.4%, 9.2% and 8.7% over a time period going from 10 to 100 milliseconds and, finally, 16.2%, 8.5% and 6.2% over a time span ranging between 100 milliseconds and 1 minute. In more absolute terms, the second graphic of each panel shows that the orders executed in less than 1 second represent 8.2%, 4.0% and 0.6% submitted on this price level for which we ultimately observe the total execution or cancellation. Considering these rates, it is impossible to rule out a potential execution intention behind the submission of short duration orders on the best price levels, no matter their actual duration. On another note, these numbers clearly demonstrate the advantage of speed when it comes to obtaining liquidity at the best available price.

From this point, in order to correctly relate these execution statistics to the short duration orders category and obtain global statistics on the orders with no execution potential, we subdivide them weather their cancellation takes place less than 10 milliseconds, between 10 and 100 milliseconds or 100 milliseconds to 1 minute after their submission. When it comes to the totally executed orders initially submitted on the second best price level, we already denote an important decrease in the proportions related to those that are subject to a rapid execution. This drop is particularly marked for the orders executed in less than 10 milliseconds which now represent 1.5%, 1.3% and 0.9% of all orders submitted on this price level that have ended in a total execution. These proportions translate into 0.07%, 0.04% and 0.01% of all totally executed and cancelled. At this point, we consider important to establish a proportion of totally executed orders with regard to those submitted on a price level that are totally executed and cancelled to act as a limit between the orders exhibiting an execution possibility and those who do not. We consider 0.1%, or 1 out of 1000 orders as a fair threshold for this task. Consequently, because of their execution rates that still show a decreasing trend beyond the second best price level of submission, we claim that the execution potential for short duration orders with a duration smaller than 10 milliseconds is virtually inexistant when submitted above the best price level. For the MDAX and SDAX stocks, the same applies to the orders with a duration between 10 and 100 milliseconds. Regarding the 10 milliseconds to 100 milliseconds duration orders submitted on the *DAX15* stocks, we shift this conclusion by one level. Indeed, Figure 5.8 shows that when it comes to the 3rd price level, 0.07% of the orders for which we observe the total execution or cancellation are executed inside a timeframe going from 10 milliseconds to 100 milliseconds, which we consider neglectable. Finally, regarding the orders executed 100 milliseconds to 1 second after their submission, the proportions cross our threshold from the 5th, 3rd, and 2nd submission price levels with respect to the *DAX15*, MDAX and SDAX components. Figure 5.8 actually shows that no matter the stock group, after these price levels, the executed orders proportions decrease and rapidly becomes way under the previously defined 0.1% threshold. To the question of knowing if our threshold is too generous, we have to keep in mind that in each case, these proportions eventually reach zero as we keep increasing in price level of submission. We conclude that over our three months period, 36.6%, 48.4% and 26.3% of short duration orders have been submitted with virtually no execution potential. Excluding the best price level, these proportions jump to 63.9%, 93.5% and 100%. These *non-executable* orders actually represent 10.6%, 16.1% and 13.3% of all successfully tracked orders, which we cautiously consider as a proxy for the total activity regarding the DAX15, MDAX and SDAX components. It is interesting to note that regarding the SDAX stocks, such orders submission and cancellation represent between 21% and 46% of all the activity observed between the 2nd and the 18th price levels.

5.3 Limit orders sequences and quote-stuffing activities

Having excluded the execution possibility as a motivation for the submission and cancellation of an important part of the short duration orders, it is now worth asking what drives some market participants to engage in such activities. In this context, we identify quote stuffing activities as a potentially valid explanation for, at least, some of these behaviors. As introduced before, quote stuffing operations consist in the submission and cancellation of several orders over very short time periods in order to clog the stock market and create confusion among investors. Widely observed by Yergeau (2018) in the Xetra auctions during the time period covered by our dataset, we consider possible for these operations to be at the source of some of the identified short duration orders.

5.3.1 Short duration orders sequences identification

Given the quote-stuffing definition, the traces left by these activities in the LOB data should take the form of sequences of orders quickly submitted and cancelled. By restricting our search to the previously identified short duration orders, we claim that the quick cancellation part is automatically covered. Consequently, in order to relate some of these orders to potential quote-stuffing operations, we have to demonstrate that their submission takes place in rapid sequences. As a first step in performing this task, we consider relevant to analyze the time intervals taking place between such orders submissions, which cumulative distribution functions are presented in Figure 5.9. We note that when it comes to *DAX15*, MDAX and SDAX components, there is no interval between short duration orders submissions in 5.97%, 7.07% and 5.25% of cases. In these situations, the involved orders submission, which can take place on the same or different price levels, have been reported through same Xetra EnBS LOB update message. For the orders whose submissions are not reported concurrently, we observe minimal interval values of 0.151, 0.127 and 0.153 microsecond. We relate the gap between the absence of interval and these values to the latency between the market participants systems and those of Xetra. We suppose that unless two or more orders submissions are simultaneously transmitted by a trader to Xetra, their impact on the LOB should be broadcasted by Xetra to all market participants through sequential messages, which may lead to these types of floor delays. Back to Figure 5.9, we observe that once past these values, the CDF values rapidly increases to show that 33.4%, 51.8% and 68.3% of the intervals between short duration limit orders are smaller than 10 microseconds. These proportions increase to 70.8%, 76.8% and 81.1% regarding the intervals smaller than 1 second.

As expected, the previous information suggests that most short duration orders do not join the book alone and we consider highly relevant to keep investigating this aspect. Therefore, we define a simple criterion that will be used to establish if whether or not, two or more short duration orders are part of a sequence. After some preliminary analysis, we decide to consider two orders as part of the same sequence if they are both submitted within a time interval smaller than 1500 microseconds (1.5 seconds). Although arbitrary, this time span allows to reduce the negative impact that a lost order can have on an entire sequence. Previous Figure 5.9 shows that for the *DAX15*, MDAX and SDAX components, 75.5%, 80.1% and 85.9% of *inter* short duration orders submission intervals are smaller than this value. We do not consider concurrently reported orders for which the submissions are neither preceded nor followed by another one inside the 1500 ms interval as composing a sequence. Finally, while orders submitted on different price levels may be part of the same sequence, we only consider orders taking place on the same book side. Based on these conditions, Table 5.2 presents the general results produced by this grouping exercise. As suggested by the previous information regarding the intervals, we first notice that only 11.3%, 9.8% and 6.1% of the short duration orders appear to take place alone. Second, we observe that 84.8%, 80.5% and 73.0% of the identified groups are composed of 2 to 9 orders. Although representing a vast majority of the identified sequences, the importance of this segment decreases when it comes to the proportions of orders involved. Indeed, their components represent 43.4%, 31.7% and 19.0% of all successfully tracked short duration orders. Nevertheless, from our point of view, these orders group actually show an important clustering phenomenon in short duration orders submission. In fact, we find preferable to identify these groups as small *clusters* rather than actual sequences. It is quite possible for some of them to result from small activity burst. Similar conclusions may be made regarding the 10 to 49 orders sequences. While second in importance regarding the number of sequences for our three stock groups, they actually come first regarding the number of involved orders for the MDAX and SDAX components with 33.6% and 36.9% of all identified short duration orders. Slightly different for the DAX15 stocks, the 33.4% of orders proportion is second. Once again, we cannot rule out the possibility for these sequences to represent orders clusters, resulting from more important peaks of activity.

Although the general concept of quote-stuffing is simple, there are no exact criteria by which it is possible to determine that an order sequence is actually related to such activities. This is especially true given the fact that we have no information on the identity of the owner of the orders and on whether they are submitted by an algorithm or not. However, from the point where at least 50 short duration orders are submitted with less than 1.5 seconds apart, we find relevant to suspect that we may be in presence of systematic and pre-programmed activities, possibly related to quote-stuffing. Nevertheless, in a preoccupation of rigor, we consider the 50 to 99 orders sequences segment as a buffer allowing to make the distinction between the previous groups potentially related to more natural activities and the sequences presenting a high potential of quote-stuffing activities. Consequently, we choose to leave aside these 12 096, 31 183 and 36 890 sequences that encompass 5.22%, 10.3% and 15.6% of all successfully tracked short duration orders for the *DAX15*, MDAX and SDAX components.

By the previous elimination process, we identify the 5 653, 16 678 and 19 152 short duration orders sequences involving 100 orders and more as those with a high potential of quote-stuffing activities. While only accounting for 0.25%, 0.78% and 1.51% of the identified sequences, they actually involve 6.7%, 14.6% and 22.5% of the short duration orders successfully tracked for the DAX15, MDAX and SDAX components over the trading period of interest. It is interesting to observe that while such sequences have been identified for all the components of our three indexes and observed at least once on every trading day, those including 500 orders and over are observed for 14, 39 and 41 out of the 15, 50 and 50 stocks in play and do not take place on all trading days for the *DAX15* and MDAX components. Providing a visual support by presenting the distribution of these sequences over time, the graphics presented in Figure 5.10 point in this direction. Indeed, they show that the SDAX sequences, which are more numerous, involve more orders and are more equally distributed over time than those observed on the *DAX15* and MDAX components. Although we have little information available to explain this phenomenon, we identify the highest proportions of orders whose track have been lost before cancellation or execution as a potential explanation for the largest number of important sequences observed on the SDAX components than on the *DAX15* or MDAX. Consequently, given the potential disadvantages associated with these orders loss, in the next section, we will focus exclusively on the sequences observed on the SDAX stocks. Given the high orders tracking success rates for these securities, we consider that the sequences have a higher potential of being neat and representative.

5.3.2 Quote-stuffing algorithms identification

When it comes to quote-stuffing, as claimed before, Egginton, Van Ness et al. (2016) have performed important works on this topic. However, because they work on the NYSE Trade and Quote (TAQ) database for the year 2010, they have to use the number of quotes updates as a proxy for the detection of quote-stuffing activity periods, which only provides the information regarding the best price levels. To some extents, such data is similar to the Xetra data processed in Chapter 1 with the exception that our original dataset contains the LOB updates up to the 20th price level. As far as we are concerned, given the data at our disposal, we expect to be able to identify quote-stuffing related orders submitted anywhere on these depth levels. Also, since the TAQ database provides no information regarding the orders submission, cancellation, and execution, they have to link the identified periods of quote-stuffing to the NASDAQ TotalView-ITCH to obtain the information regarding these elements. As claimed before, using the methodology developed in Chapter 3, we are able to obtain all the information regarding the life-cycle of several orders, particularly on the SDAX components, which will represent the main focus of this section. Finally, Ness et al. (2016) have to divide each trading day into 1 minute periods in order to detect quote-stuffing activities, which we do not have to do. Indeed, the methodology presented in the previous section allows us to establish the almost exact beginning and end of each identified orders sequence, whether it has a duration of a few milliseconds or several minutes. All these elements make us well suited to identify recurrent algorithmic signatures and to analyze different characteristics of the orders composing them.

General example

To present some of the elements that we consider important in order to analyze the behavior of the algorithms responsible for some of the short duration orders sequences identified, we begin by the presentation of our potentially simplest identified sequence. Table 5.3 provides a chronological list regarding some of the 100 orders composing a sequence that have been observed on the ask book side of the KWS stock over the 1.9 seconds time period going from 15:05:01.362 to 15:05:03.316 on 2013-04-26. As shown in this table, each of the 100 limit orders consists in the same 14 shares quantity offered at a 277.05 EUR physical price. These orders average duration is 9.36 milliseconds with a standard deviation of 0.12 millisecond. At this point, it becomes obvious that such small and similar durations may not be achieved other than by an algorithm through which the orders submission and cancellations are systematic and pre-programmed. Indeed, these orders have been submitted at a rate of 51.49 events by second, which is clearly out of the actual human limits and even out of those of the firing power of modern machine guns.

By visually presenting our sequence example, Figure 5.11 (xxx) supplements the data presented in Table 5.3 regarding its global characteristics and the order book environment in which it takes place. In Panel A, we present the complete sequence, in Panel B, we present the first 200 milliseconds in order to have a closer look in a kind of zoom, which is once again amplified in Panel C where we present 40 milliseconds. In these panels, as well as the other examples that will be presented in this section, the wide black line represents the limit orders identified as part of the sequence, which take place on the ask book side in this example. The thin light grey, medium grey and dark grey lines indicate the prices corresponding to the best, second best and third best price levels of the book side affected by the sequence at any time. This is the reason for which the wide black line is always horizontally separated by a thin grey line whose level of darkness indicate the price level number of each order in sequence. It is important not to confuse this representation with two concurrent orders who would take place on two different physical prices at the same time. Finally, in Panel A, the thicker light grey line indicates the opposite book side best price level. In this case, this line shows that the best bid price have remained constant at 275.35 EUR for the entire duration of the sequence, which, as we will see later, is generally the case during an algorithmic sequence taking place on one side of the book. Back to the sequence itself, as already pointed out, these figures first show that all the orders have been submitted on the same physical price, i.e. 277.05 EUR. Second, as also presented in Table 5.3 *Submission level - Number* and *Context* columns, it is possible to observe that all of them are submitted inside the bid-ask spread, each time creating a new best ask price level. Easily observed in Panel B and C, since each order remains the only one standing on this price level between its submission and cancellation, after each cancellation, the best price moves back to the 277.10 EUR physical price. The same applies to the second best ask price level that becomes 277.10 EUR during the life of each order in sequence while returning to 277.35 EUR after each cancellation. Finally, we observe a time span between each order cancellation and the submission of the next one. Detailed in the *Time since previous order cancellation* column of Table 5.3, these periods have a duration average of 10.3 milliseconds and a standard-deviation of 0.09 millisecond. Once again, these particularly similar small values argue in favor of the pre-programmed character of the sequence.

Classification methodology

Because they take place in the time-price space, algorithmic sequences of orders present different characteristics both general and related to their constituent orders. In this context, there are several angles from which it might be possible to analyze them. Therefore, we choose to characterize and group the algorithmic sequences based on their physical prices of operation, the depth levels relative to the best price on which the orders are submitted and, the involved orders continuity in time. As presented in Table 5.4, by grouping orders sequences based on this information, we are able to establish a list of algorithmic signatures, or profiles. We therefore use this table as a frame of reference for the detailed analysis of algorithmic behaviors throughout this section. For each profile, the *Ranked physical prices proportions* presents the proportions of orders corresponding to the most frequent physical prices on which orders are submitted during a sequence. As examples, a 100% proportion for the 1st price indicates that all the orders are submitted on the same physical price and a 67% 1st - 33% 2nd prices distribution indicates that two-third of the orders are submitted on one physical price and one third on another price. As presented in this table, we analyze algorithmic signatures for which all the orders are distributed on up to ten physical prices. We consider that this limit provides an interesting scope to the study, despite the fact that several sequences consist in orders taking place on more than ten prices. In the *Continuity* section, we provide the proportions of sequences orders that we consider as *Non-Contiguous* (NC), *End-to-end* (ETE), and *Overlapping* (OLP). We consider an order as Non-Contiguous if there is a time gap between its submission and the cancellation of the preceding order in sequence. We consider an order as End-to-end if its submission time exactly matches the preceding order cancellation time. As a last category, we consider an order as Overlapping if it is submitted while the preceding order in sequence is still present on the order book. Finally, the *Orders submission price levels* section presents the proportions of orders in sequence by relative depth level of submission. For each of the three best relative price level, we make the distinction between the orders whose arrival creates a new price level and those who are submitted on an already existing price level. The ≥ *4* column reports the proportions of orders submitted deeper than the third best price level, regardless of whether they take place on a new or an existing price level. It is important to note that since it is almost impossible to find identical proportions among the orders sequence, we have rounded all the proportions discussed to the nearest 3.33% value. It should also be noted that we have excluded the algorithmic profiles grouping five sequences or less in order to keep the analysis in reasonable proportions. For each algorithmic signature, Table 5.4 also present a Type, which is an arbitrary sequential number which we produce so that the first digit matches the number of main physical prices over which the sequence takes place. The column *Count* provides the number of orders sequences represented by the profile, the column *S* corresponds to the number of different stocks on which it has been observed and the column *TD*, the number of different trading days. Finally, we provide the proportions of sequences taking place on Bid and Ask book side, the minimum and maximum number of orders involved in a sequence, the minimum and maximum sequence duration in *minutes:seconds.deciseconds* as well as the minimum and maximum orders arrival rates in orders by second. Table 5.5 provides a sequence example for each algorithmic signature presented in Table 5.4. Several of these examples will be detailed bellow.

Algorithmic profiles

Table 5.4 algorithmic profiles 1.1 to 1.3 involves the submission of orders on a single physical price. It is interesting to note that overall, the number of sequences presenting this idiosyncrasy is very small, even when considering sequences who are not characterized by the algorithm signatures presented in the table. Indeed, we observe that only 63 out of the 19 152 identified sequences involving at least 100 orders in which more than 80% of orders are submitted on the same physical price. The example presented before as a general sequence introduction (Table 5.3 and Figure 5.11) belongs to the algorithmic Profile 1.1. Indeed, 100% of its orders have the same physical price, 100% are submitted on a new best price level and they are all separated by a time space, which makes them 100% non-contiguous (NC). Signature 1.2 sequences present the same characteristics except that the orders are submitted on a new second best price level. Table 5.5 Example 1.2 is identified as part of this category. In this case, the sequence consists in 122 bid orders with a price of 4.165 EUR. As presented in Table 5.6 and visually shown in Figure 5.12, these orders have an average duration of 137 milliseconds and the interval between a cancellation and the next submission is way narrower with an average of 1.06 millisecond. As presented in Panel C, during these periods, the second best price level moves back to its original 4.164 EUR physical price. This specific example is very interesting in terms of underlying goal behind the strategy. Indeed, assuming that the owner of these orders would actually be interested in buying two stocks for 4.164 EUR each during the period, why wouldn't he submit a single order for the complete 16.9 seconds duration instead of proceeding in 122 submissions and cancellations? From our point of view, the absence of rational answer to this question points out to the possibility that the investor may have had no real trading intention, especially given the submission of the second best price level for the orders submissions and the tiny number of shares characterizing the orders. Therefore, quote-stuffing becomes a logical explanation for this behavior. Back to Table 5.4, it is possible to observe that Signature 1.3 is similar to 1.1 and 1.2 with the difference that the orders are submitted on a new third best price level, which eliminates the need for an additional example.

Algorithmic profiles 2.1.1 to 2.13 involve the submission of orders on two main physical prices. As reflected in Table 5.4 we observe that overall, at least 23% of the identified sequences belongs to this global category, which is definitely the most important when performing a classification based on this characteristic. Profiles 2.1.1 to 2.11 refer to algorithms using the same physical price distribution in their operations. Indeed, the order prices are distributed among the two most frequent prices in a two-thirds - one-third way. This observation becomes obvious when looking at Example 2.1.1 whose orders are detailed in Table 5.6 and visually presented in Figure 5.12. As shown in Panels B and C of the figure, the actual sequence consists in two 16.30 EUR orders followed by one 16.295 EUR order. In the Xetra 10 to 49.995 EUR range price for a stock, the tick size is 0.005 EUR, which corresponds to the difference between the two physical prices involved in the sequence. Unlike the previous examples, we observe that there is no time gap between the cancellation and the submission of an order in this example, which lead to the 100% end-to-end (ETE) characteristic of the algorithmic profile 2.1. We have to indicate that in Figure 5.12 Example 2.1 Panels B and C, the thin white line between the 16.30 EUR orders are added by us to point the change in order quantity that goes from 81 to 334 shares. It does not represent a real time span as observed in the previous example. In fact, despite impossible to prove, because of this absence of delay between each order, we believe that these sequences could actually consist of a single order being modified 101 times instead of multiple submitted than cancelled orders. Indeed, no matter the performances of the involved systems, cancelling than submitting a new order would probably leave at least a small time span between each operation. From our point of view, the quantity of shares characterizing the orders may support this theory. Indeed, in Table 5.6, we notice that the 16.295 EUR orders consisting in 81 shares are followed by a 16.30 EUR order that also consist in 81 shares, which could be explained by an order price modification. In each case, the 16.30 EUR order quantity afterward change to 334, which could also be related to a simple order quantity of shares modification. Finally, the 16.30 EUR order consisting in 334 shares could, once again, be modified to become a 16.295 EUR order representing 81 shares. We have to acknowledge that in this last case, the fact that two order characteristics would have to be concurrently modified slightly weaken our theory but, since still highly possible, it does not discard it. We have to notice that in this example, all the operations take place within a 829 milliseconds time period which, from our point of view, definitively rules out the possibility of human actions.

Algorithmic profiles 2.1.2 to 2.1.13 all describe situations similar to that of signature 2.1.1 with some small differences. We actually consider these differences a collateral effects of our methodology in which we round the proportions to obtain much clearer groups. Figure 5.12 presents an example of profile 2.1.13 algorithmic sequence, which physical prices distribution presents the more important level of discrepancy with regard to 2.1.1. From our point of view, this particular example shows that although operating on a very short timeframe, these algorithms seems to be adaptable. Indeed, after observing a change in the out-of-sequence best ask price level after about 37.29 milliseconds in sequence, the algorithm appears to react by modifying the physical prices of the orders submitted through the remaining of the sequence. The first order presenting this type of modification is observed less than 1 millisecond after the 8.115 EUR pre-sequence best price level disappearance. This also shows that the orders position relative to the best price may actually be more important than the physical price itself. Leaving aside the first 8.114 EUR 452 shares order that appears slightly different, Table 5.6 shows that the actual sequence pattern begins with one 8.114 EUR 48 shares order followed by one 8.114 EUR 410 shares order, then, by an 8.113 EUR 508 shares orders. From the time where the new 8.12 EUR out-of-sequence best ask price level seems considered by the algorithm, the previous 8.114 EUR 48 shares orders become submitted on the 8.119 EUR physical price with the same quantity of shares, the 8.114 EUR 410 shares orders are replaced by 8.119 EUR 472 shares orders and finally, the 8.113 EUR 508 shares orders are substituted by 8.118 EUR 890 shares orders. As before, these orders stand 0.001 EUR and 0.002 EUR under the new out-of-sequence best price which, for the 0 to 9.99 EUR stock price range, correspond to respectively 1 and 2 ticks. This modified pattern goes on until the end of the sequence. We are unable to find the rationale behind the change in the number of shares for two out of the three orders in sequence. However, these modifications may suggest that the involved algorithms have their own internal mechanisms to determine the number of shares to include in each order.

As shown in Table 5.4, algorithmic profiles 2.2 to 2.11 have the same characteristics than 2.1.1 in the fact that they operate in a two-thirds – one-third mode in terms of physical price and that 100% of the orders are end-to-end. In these cases, the differences lie in the orders submission patterns with regard to the best prices. Figure 5.12 and Tables 5.5 and 5.6 Example 2.2 is highly representative of this situation. It is easy to retrieve all the characteristics of the previous signature 2.1.1 except that the orders submissions (or modifications) are performed around two newly created second best ask prices at 75.47 EUR and 75.48 EUR. As presented through Example 2.4, signature 2.4 presents another situation in which the original best ask price level lies between the two physical prices favorized by the algorithm. As shown in Table 5.6, a 7.069 EUR 127 shares is first submitted on a new second best price level. Second, still based on our previous theory, the order quantity of shares appears increased to 700 shares. Since performed on the same price, we consider this operation as a new order submitted on the existing second best ask price level. Finally, the order seems modified to become a 7.059 EUR, once again, 127 shares order, which operation creates a new best ask price level. Repeated 40 times, this sequence results in one third of orders creating a new best price, one third creating a new second best price and one third of orders replacing the single order present on the same second best price. Without requiring examples since very similar to 2.4, profiles 2.5 to 2.9 consist in different combinations of analogous operations. When it comes to algorithmic signatures 2.10, 67% of the involved orders affect the fourth best price level and more. Similarly, in the 2.11 case, all the orders take place on such deep price levels. As shown before, the execution potential for even a single order involved in such sequence becomes virtually inexistant.

Before moving forward with the next algorithmic profiles, we consider relevant to come back to some general characteristics of signatures 2.1.1 to 2.11. We consider important to point out the fact that these widespread sequences both in terms of number of stocks for which they are observed and in terms of trading days, only take place on the ask side of the order book. Indeed, we find difficult to explain that unlike the other profiles that are observed on both sides of the book, these signatures are never observed on the buy side. We also consider important to notice the incredible arrival speed regarding the orders involved in these sequences. Indeed, as presented in Table 5.4, the orders arrival rates generated by these algorithms range between 33 and 1406 orders arrival by second. No matter that the orders are actually modified or cancelled during the sequences, we have to double these numbers to take the orders termination into account and obtain an actual event arrival rate. Resulting from these high rates, none of these sequences, which involve between 100 and 245 orders, ever take place over a time period longer than 3.7 seconds. Still focusing on our detailed algorithm analysis objective, we find important to mention that in these particular cases, a one minute time period activity aggregation could have led to an important under-estimation of these activity peaks events arrival rates. Table 5.5 Example 2.1.4 is a good illustration for this situation. It represents a 121 orders sequence observed for the EV4 stock on 2013-03-07 around 15:45:19. Overall, we observe, a total of 132 ask orders submissions one minute time period going from 15:45:00 to 15:46:00 over which the sequence have taken place. Therefore, 121 orders are involved in the sequence and 11 orders are out of sequences. Therefore, aggregating these 132 orders over the 1 minute period, would lead to an arrival rate of 2.2 orders by second. However, during the 129 milliseconds period over which algorithmic profile 2.1.4 have actually been in operation, this rate has been established to 939 orders by second, which show the potential importance of such rates under estimations. Finally, despite obvious, we consider important to notice that by only considering the orders affecting the best price levels, it would actually be impossible to identify the algorithmic signatures 2.2, 2.3, 2.6, 2.9, 2.10, 2.11 as well as most orders taking place through profiles 2.4, 2.5, 2.7 and 2.8.

Back to Table 5.4, we now focus on Algorithmic signatures 2.12.1 and 2.13 which, despite also operating on two main physical prices, suggest a very different behavior from what has been observed in 2.1.1 to 2.11. In general terms, it is possible to observe that the sequences generated by these profiles often consist in more orders which, because they are scattered over longer time periods, lead to smaller arrival rates. Indeed, we observe that the average number of orders involved in these sequences is 323 with maximum reaching more than 2000 orders for profiles 2.12.1, 2.12.5, 2.12.9 and 2.12.10. The overall average duration for these sequences is 3 minutes 2 seconds and this duration reaches more than 10 minutes in 23 observed cases. With a single exception where the arrival rate reaches 256 orders by second for profile 2.12.1, all the sequences related to these signatures have an arrival rate lying between 1.4 and 2.1 orders by second, which is relatively low with regard to the previously analyzed profiles.

Although it involves only 126 orders, which is below the previously established average, Example 2.12.1 presented in Tables 5.5, 5.6 and Figure 5.12 is highly representative of the algorithmic profiles 2.12.1 to 2.12.16. As shown Panels A and B of Figure 5.12, the algorithms behind such signature operate by submitting two almost concurrent short duration orders on two different physical prices. Once these two orders cancelled, nothing occurs during a time span way longer than what we have observed with previous profiles. Then, a similar two orders submissions and cancellations pattern is performed again and so on. As seen in Table 5.4, no matter the sub-profile, almost all orders are submitted on a new best price level. In the particular Example 2.12.1, because of the graphics scale, it is impossible to notice this behavior in the Panels A and B. However, with a closer look, Panel C provides a good visual representation. It is possible to note that the 1.718 EUR shares are submitted slightly before their 1.717 EUR companions. Supplementing this information with Table 5.6 Example 2.12.1 *Time Since Previous Order Submission (TSPOS)* column for the 1.171 EUR orders, it is possible to establish that the average time between the two orders submissions is 3.02 milliseconds. As shown in Panel C, during this short time period, the 1.718 EUR order stands on the best ask price level. Given the usual market mechanism, the 1.717 EUR order submission causes the 1.718 EUR order to automatically fall on the second best price level. As shown in Panel C, the 1.718 EUR order is later cancelled while standing on this price level, because of the 1.717 EUR order still present on the best price level. We note that the 1.717 EUR orders cancellation occurs with an average delay of 2.61 milliseconds after that of the 1.718 EUR orders. From an order continuity point of view, the negative values in Table 5.6 Example 2.12.1 *Time Since Previous Orders Cancellation* (TSPOC) for the 1.717 EUR orders shows that on average, these orders are submitted 14.87 milliseconds before the 1.718 EUR orders cancellation. In this context, we consider these orders as overlapping (OLP). On the other hand, the same TSPOC column values for the 1.718 EUR show that on average, these orders are submitted 1.258 seconds after the 1.717 EUR orders cancellation, which make us consider them as non-contiguous (NC). Consequently, the fact that the number of 1.717 EUR and 1.718 EUR orders are almost the same leads to the 50% NC – 50% OLP information characterizing the 2.12.1 to 2.12.16 algorithmic signatures in Table 5.4. In this particular example, the 1.258 seconds average time between each two orders set explains the relatively low order arrival rate of 1.6 orders by second for this sequence. With their own orders duration and inter-order time spans, this observation can be generalized to all the other sequences resulting from similar algorithmic signatures. Table 5.6 Example 2.12.1 shows an interesting algorithmic behavior when it comes to the quantity of shares represented by each order. Indeed, regarding the 1.717 EUR orders, this quantity is 3028 for the first twenty-two orders, 3539 for the next twenty-four and 3191 for the last seventeen. When it comes to the 1.718 EUR orders, this quantity remains the same at 3167 shares among the sixty-three of them.

As seen before, algorithmic signature 2.13 is similar to 2.12.1 with the difference that the two concurrent orders are separated by the pre-sequence best price level for the affected book side. This leads to the identification of 50% of orders taking place on a new best price level (Level 1) and 50% on a new second best price level (Level 2). Once again, 50% of the orders are considered as non-contiguous (NC) and 50% as overlapping (OLP). Figure 5.12 Example 2.13 presents this signature, which as shown in Panels A to C, is very similar to 2.12.1. Panel B focuses on the submission of the two recurrent orders composing the sequence. We observe that at first, a 12.555 EUR order is submitted above the original best ask price level, which creates a new second best price level. Then, a 12.545 EUR order is submitted below the original best price level, which causes the still present 12.555 EUR order to automatically fall on the third best price level. As presented in Panel C, this order still belongs to this price level when its cancellation occurs, which is followed by the cancellation of the 12.545 EUR order that makes the best ask price level to move back to its original 12.55 EUR physical price. As before, while all the orders with the highest price consist in the same number of shares (800), we observe an evolution in this quantity for the orders with the lowest price which changes seven times during the sequence.

Patterns similarities

Having performed a detailed analysis of the algorithmic profiles involving orders mostly takin place on two physical prices, it now becomes interesting to see how the different observed behaviors are found through the remaining signatures presented in Table 5.4 in which the number of involved physical prices is increasing. Beginning with profile 3.1, it is possible to denote important similarities with previous signature 2.1.1 despite the obvious exception of the addition of a third physical price over which orders are present. Indeed, 100% of orders are considered as end-to-end which, once again, suggests the possibility of orders modifications instead of cancellations and submissions instructions. We also note that 60% of the orders are submitted on a new best price level and 40% are submitted on the new price level created by a preceding order, which at the time, is considered as an existing price level. Figure 5.12 Example 3.1.1 Panel B shows that these proportions actually hide a 2-2-1 orders pattern. Indeed, in this example, a 3.662 EUR 344 shares order is first submitted, followed by a change in quantity of orders that becomes 903 shares. In a second step, we observe a 3.661 EUR 344 shares orders taking place after the disappearance of the previous order. This 3.661 EUR quantity of shares evolves to 893 before its disappearance. Finally, the third step consists in the arrival of a 3.66 EUR 344 shares orders which eventually disappear to be replaced by the first described 3.662 EUR order with the same quantity of shares. As shown in Panel A, this *stair* pattern in which the orders appear to *descend* into the bid-ask spread is then repeated several times. As before, we observe that algorithmic profiles 3.2 and 3.3 leave the same traces as 3.1, except that they operate around newly created second and third best price levels.

Moving downward in Table 5.4, it is possible to observe that algorithmic signature 4.1.1 and 4.1.2 also present very similar characteristics to 3.1, once again, with the exception of a fourth physical price involved. Indeed, as presented in Figure 5.12 Example 4.1.1 Panels A and B, in this case, the recurrent stair pattern takes a 2-2-2-1 shape in which each cycle appears to consist in seven orders modification. At this point, we consider important to observe that as shown in Table 5.4, algorithmic signatures 3.1 to 3.3 as well as 4.1.1 and 4.1.2 all take place at very high speed in terms of orders by second. They also take place over very short time periods going from 0.1 second to 1.5 seconds.

Additionally, we have to note that none of these profiles traces have been observed on the bid side during our three months data period. These similarities with the 2.1.1 to 2.11 signatures make us believe that we could be in the presence of several versions of the same algorithm or even of a single algorithm whose parameters could vary according to the situation or the preferences of the market participants operating it. Indeed, behind the differences in terms of physical prices proportions and produced orders price level with regard to the best ask price, from our point of view, the operation mode remains almost identical.

Along the same line, Figure 5.12 Example 3.4 shows that algorithmic profile 3.4 presents several similarities with the previous signature 2.12.1. As globally presented in Panels A and B, this profile pattern first results in three grouped orders submitted and cancelled over a very short time frame. The last order cancellation is followed by a relatively long gap before the submission of the next order. In the current example, while the three recurrent orders present 11.47, 11.38 and 0.86 millisecond average durations, the average length of the gap is 1.35 second. As presented in Panel C, the distinction with profile 2.12.1 regarding the orders submission and cancellation pattern is the addition of a third order in each recurrent sub-sequence. Indeed, using example 3.4 prices, we can see that the 7.118 EUR and the 7.117 EUR orders present the same behaviors in terms of continuity and relative price level than those observed in 2.12.1. However, immediately following the 7.117 EUR order disappearance, a very short duration 7.119 EUR order is submitted, then cancelled. In fact, there is no time span between the 7.117 EUR order cancellation and the arrival of the 7.119 EUR one, which is considered as end-to-end in our orders continuity classification, leading to the 33% non-contiguous, 33% end-to-end and 33% overlapping profile for the algorithmic signature. Based on these similarities, we consider once again possible that the same algorithm family could be involved in profiles 2.12.1 to 2.13 and 3.4.

We complete the analysis of algorithmic signatures involving three main physical prices with profile 3.5, which, as shown in Figure 5.12 Example 3.5, presents an interesting pattern that we have not seen before. Indeed, as globally presented in Panel A and detailed in Panels B and C, the limit orders take place on the three existing best price levels. In this particular case, we observe that the order in play first takes place on the third best price level at 27 EUR. In a second step, its price becomes 29.995 EUR, which makes it fall on the second best price level. In a final third change, it becomes part of the existing best ask price level at 26.95 EUR. Since the 643 end-to-end orders include exactly 528 shares, once again, we believe that the sequence is made of a single order that is modified at each step. Figure 5.12 Example 3.6 shows that algorithmic profile 3.6 presents the same behavior as 3.5 with the difference that the third order in the pattern falls on a new best price level. Once again, the orders are end-to-end and present the same 387 quantity of shares. From Table 5.4, it is interesting to observe that the algorithm presenting these signatures are relatively slow with 3 to 6 orders by second. However, they are observed over time periods going from 16.5 seconds to more than 6 minutes and we have related up to 1305 orders to their operations.

Algorithms versions

Despite their different numbers of physical price of operation, profiles 5.1 to 10.1 allow us to find important similarities among the resulting orders sequences which, from our point of view, add plausibility to our theory that certain algorithms are available in several versions or that they are actually configurable. We first observe that profiles 5.1, 7.1, 9.1 and 10.1 produce patterns that could be produces by algorithms belonging to the same *family*. Figure 5.12 Example 5.1 Panel A shows that when it comes to algorithms submitting limit orders on five physical price and more, the situation becomes such that the order book is simply flooded with submissions, cancellations, and modifications of short duration orders. In this specific example, Panel B shows an actual pattern in which two concurrent orders are followed by two other concurrent orders then by a single order, all descending in price toward the opposite side best price (bid). Panels C, D and E show a small lag in the submission of the concurrent orders. Given this time span, as indicated in Table 5.5, all the orders initially take place on a newly created best ask price level.

In terms of orders continuity, our classification system considers that the 34.40 EUR, 34.395 EUR, 34.395 EUR, 34.39 EUR and 34.385 EUR are submitted concurrently while the 38.405 EUR order is end-to-end with regard to the 34.385 EUR order. This leads to the 80% overlapping – 20% end-to-end classification presented in Table 5.5. This classification is explained by the fact that, as shown before with the column *Time Since Previous Order Cancellation* (TSPOC), our cataloguing system looks at the previously submitted limit order cancellation to determine the continuity of a newly arriving limit order. However, with a very close look at Panels C, D and E, or at the detailed orders data, it is possible to note the absence of time delay between the cancellation of the 34.385 EUR order and the arrival of the 38.405 EUR order. The same applies to the 38.405 EUR order cancellation and 38.395 EUR order submission as well as for the 38.395 EUR order cancellation and, at the end of the round trip, the 38.385 EUR order arrival. Once again, this suggests that a single order could be modified instead of several orders submitted and cancelled. The same applies to the 34.40 EUR and 34.39 EUR orders for which the cancellations and submissions appear end-to-end. However, in this last case, since they take place alone, the arrival of the 34.40 EUR orders and the disappearance of the 34.39 EUR orders have to be related to actual submissions and cancellations. As claimed before, Figure 5.12 Examples 7.1, 9.1 and 10.1 show that Table 5.4 algorithmic profiles 7.1, 9.1 and 10.1 result in orders sequences very similar to profile 5.1. Indeed, with the exception of orders takin place on more physical prices, it is possible to observe a common general recurrent pattern. When it comes to profiles 6.1.1 and 6.1.2, despite also similar, Figure 5.12 Example 6.1.1 shows that the order with the highest physical price (5.969 EUR) takes place concurrently on the book with the lowest physical price order (5.864 EUR), which leads to a difference in the orders submission price levels proportions that become 83% and 17% on new and existing best price levels. As seen before, signature 6.2 exhibits the same characteristics except that the orders are submitted around new and existing second best price levels.

Closing the analysis of algorithmic signatures presented in Table 5.4, as shown in Figure 5.12 Examples 6.3 and 8.1, algorithmic profiles 6.3 and 8.1 produce very similar results despite the fact that one takes place on six physical price levels and the other, on eight. At this point, we can only appreciate the complexity of the algorithms behind such orders sequences. As shown in Panel A of Example 6.3, the orders submission, cancellation, and modifications patterns becomes very interesting and particularly difficult to describe. Example 6.3 Panel B and Example 8.1 show that despite the fact that the 8.1 example takes place on the bid side and consists in more orders, the similarities remain present. Therefore, these examples show the ability of a same algorithm, or algorithm family, to operate on both sides of the order book.

Long sequences

Finally, in order to show the proportions that quote operations can take, we consider relevant to present two examples of long, short duration orders sequences whose classification goes obviously beyond the limits of our previous methodology. First, Figure 5.13 Example 1 presents an algorithmic sequence that took place on the MLP stock on March 13th, 2013. It lasted 16 minutes and 57 seconds over which 4093 short duration ask limit orders have been observed. As often, a vast majority (92.89%) of these orders have arrived on a new best ask price level. Also, 97.53% of the orders are considered as overlapping, which have ensured their cancellation on the second best price level. Despite appearing very complex at first, Panel A shows that this sequences splits into 18 segments where the algorithm has operated from a new out-of-sequence best price level that have generally become the third best price level once the orders in sequence considered. At this point, the question of whether the algorithm itself produced these out-of-sequence best price changes remains open. As shown in Panel B, which presents a closer look to one of these segments, it is possible to recognize one of the previously exposed patterns. Indeed, the stair shape of the orders modifications processes and the fact that these orders appear in pairs is very similar to what we have observed with profiles 5.1,7.1, 9.1 and 10.1. We have to mention that this time, the recurrent pattern takes place on up to 20 physical prices. From our point of view, this once again shows that several quote-stuffing algorithms leave similar traces, despite the fact that the complexity and extent of their operations may widely differ. Figure 5.13 Example 2 presents a second example of long algorithmic sequence which have also been observed on the MLP stock, this time on 2013-02-18. It consists in 2507 orders over a 9 minutes 7 seconds period. Also appearing complex when looking globally at Panel A, it is however possible to note that it actually consists in 24 orders sub-sequences, each starting from a changing best ask price support. As shown in Panel B, they are in fact series of concurrent and single orders, which in terms of continuity show 27% overlapping and 73% end-to-end orders proportions. Once again, 100% of the orders take place on a newly created best ask price level. From our point of view, the most interesting fact regarding these orders is that in their stair shaped arrival and disappearance patterns, they consecutively take place on up to 97 physical prices. As for several previously presented sequences, this is made possible by a wide bid-ask spread that is divided into 0.001 EUR ticks given the price range of the orders.

General comments

Given the particular regularity in the patterns and the order characteristics observed before, it appears obvious to us that each of the analyzed sequences have been produced by the operations of a single algorithm. At this point, any doubts about the quote-stuffing objectives linked to algorithms producing this type of order sequences are dispelled, mostly by lack of any other plausible explanation. Although individual strategies may vary, we consider very likely that these algorithms intent to overwhelm the other market participants trading systems with significant flows of orders submissions, modifications, and cancellations. This becomes particularly clear when we analyze the situation from our general high-frequency data exploitation point of view. Indeed, in the specific case of Xetra, in order to keep their order book up to date for a given stock, any real-time customer of the Xetra Enhanced Broadcast Solution system must consider all operations performed on this stock at any time. We have seen in Chapter 1 that depending on the context, it is possible that several operations are required in order to take into account the submission or cancellation of a single order. Therefore, when these operations are made necessary up to several hundred times by second by a quote-stuffing algorithm, we can easily figure out that the systems of numerous market participants can get congested and slowed down. We actually believe that small investors with more limited technological means may be more affected by this type of strategies. It is also evident that the complexity and the extent of the used patterns can be the source of much confusion. Indeed, by putting ourselves in the shoes of a traditional trader interested in a stock subject to one of the two long algorithmic sequences presented in Figure 5.13, the question that comes to mind is *"What is going on with this stock and how long will it last again?"*.

5.4 Liquidity or activity?

As a last analysis presented in this chapter, we consider relevant to focus on the liquidity provided by the short, medium, and long duration orders with regard to the total activity that they generate. In order to quantify the liquidity provided by these three orders categories, we define a simple measure based on some of their characteristics. It is simply a question of defining a standardized unit of temporal liquidity which will become a basis of comparison for all orders. Since it is obvious that an order composed of a hundred shares with a 10 EUR price does not provide the same quantity of liquidity if it stays on the book for 1 minute than an order composed of a hundred 100 EUR shares, we cannot use the sole quantity of shares to perform this task. Consequently, as a liquidity provided measure, we use the total EUR value of each order. In order to consider the temporal aspect of the provided liquidity, which is the main aspect of our analysis, we integrate this value over the duration time of the order. Thus, we use the following formula to obtain the temporal liquidity provided by any order $i \in \mathbb{N}$ whose arrival and cancellation time correspond to $t_i^a \in \mathbb{N}$ and $t_i^c \in \mathbb{N}$ such that $t_i^a < t_i^c$:

(5.1)	$$\begin{aligned} liquidity_i &= \int_{t_i^a}^{t_i^c} (price_i * quantity_i) dt \\ &= price_i * quantity_i * duration_i \end{aligned}$$

where, $price_i$ corresponds to the order price, $quantity_i$, its quantity of shares, and $duration_i$, to its time between submission and cancellation. Although not important for relative analysis purpose, we choose to express the duration in minutes, which will later be interesting for absolute data interpretation.

For each order category : short, medium, and long duration, Figure 5.14 presents the proportion of liquidity provided obtained using equation (5.1), with regard to the proportion of activity generated, in terms of orders submitted. We only focus on the orders submitted on price levels 1 to 10 in order to avoid the potential issues related to the orders whose track is lost because they exit the 20 visible price levels window. This explains the slight proportion differences between these numbers and those presented before. Panels

A, B and C, presents the information relative to the usual *DAX15*, MDAX and SDAX components. Inside each panel, the small circles represents theses proportions for each individual stocks and the large circle for entire stock group. Across the three panels, the most obvious fact is that no matter the proportion of activity generated by the short duration orders, they provide almost no liquidity in relative terms. Indeed, for the DAX15, MDAX and SDAX components, while accounting for 33.5%, 37.0% and 51.8% of the successfully tracked orders, they only account for 0.40%, 0.07% and 0.21% of the liquidity provided which, from our point of view, qualify as marginal. On the other hand, while representing only 6.7%, 14.1% and 5.8% of the activity, the long duration orders provide 62.5%, 86.9% and 86.9% of the order books liquidity. With 37.1%, 13.0% and 12.9% of the provided liquidity, although they are not great providers of liquidity, medium duration orders stand in the middle with liquidity provided to generated activity ratios closer to 1. These data clearly show us the importance of long duration orders, which are less numerous, with regard to the liquidity present in the order books over time.

In order to put this information into perspective, we find interesting to compare the liquidity provided by the short duration orders observed on the SDAX index components to an average order. To perform this task, we simply standardize the liquidity using the average *Quantity of shares * Price* value for all SDAX orders successfully tracked from submission to cancellation over our three-month data period. Based on formula (5.1), this corresponds to an average SDAX order EUR value. To minimize the impact of the orders whose track is lost because they move above the 20th price level, we only consider the orders submitted on depth level 1 to 10 to obtain this quantity. For the 29.6 million of orders meeting these criteria that belong to the SDAX components, this value is 9 272.65 EUR. Based on equation (5.1), for any order $i \in \mathbb{N}$, the standardized liquidity is provided by the following expression :

(5.2)	$standardizedLiquidity_i = \frac{liquidity_i}{quantity * price * 1\,minute}$

where $\overline{quantity * price * 1\, minute}$ obviously relates to the previously described average order value (9 272.65 EUR), assuming that such order stands on the book for one minute. Although appearing somewhat arbitrary, we consider this expression interesting for interpretation purposes. Indeed, no matter its duration, price, or quantity of share, it becomes possible to claim that an order $i \in \mathbb{N}$ for which $liquidity_i = 1$ have provided as much liquidity as an average value order over a one minute period. Table 5.7 presents this standardized liquidity measure for the short duration orders of the 25 SDAX stocks for which the greatest number of them was observed over our three-month data period. Once again, we split the short duration orders into three categories : smaller than 10 milliseconds, between 10 and 100 milliseconds and between 100 milliseconds and 1 second. For each stock, we present the number of such orders as well their aggregated standardized liquidity measure. It is then very interesting to see small quantity of liquidity provided by the orders whose duration is smaller than 10 milliseconds. Indeed, as an example, we note that the 544 013 orders meeting this criterion for the SFQ stock have provided no more liquidity than our average order for a 13 minutes 54 seconds period over the 61 trading days. Despite the fact that the standardized liquidity appears more important when it comes to the 100 ms to 1 min duration orders, these data show that overall, with regard to SDAX securities, short duration orders generate a lot of activity compared to the liquidity provided.

As expressed by Ness et al. (2016), market participants criticize quote-stuffing operations, claiming that they create a *false sense of the true supply and demand for a stock*. Based on the elements presented in this section, we actually consider possible to extend this critic to all short duration orders. Indeed, we have shown that the proportion of liquidity provided by these orders is simply inadequate with regard to the generated activity in terms of orders submission, cancellations, and modifications. We actually consider that they may lead to a suboptimal resources usage, which should particularly affect the smallest market participants whose technological infrastructures may be more limited. In addition to the real-time processing issues discussed before, we also have to consider their impact on everything that is done outside the market such as, analysis, models estimation and trading strategies testing. By aggregating the numbers obtained over this chapter, we note that 39% of all the orders are cancelled less than 1 second after their submission, which we have described as short duration. We can therefore expect that a similar proportion of the data in terms of quantity and physical size should be directly attributed to these activities. As presented in Chapter 1, the size of the data resulting from a single trading day on Xetra, which was only the 13th most important stock market in the world in 2019, is counted in gigabytes. Based on their proportion, we can establish that it is quite possible that the data resulting from the short duration orders also daily settle in gigabytes, which quickly become terabytes when accumulated for analysis and future usage. Although data storage is not considered as a problem anymore, the fact remains that processing large amounts of data still presents certain challenges, which, from our point of view, may be unnecessary amplified by these orders that do not represent a true source of liquidity. Here again, it would be very interesting to analyze the effects of the Order-to-trade ratios and the Fees for excessive system usage enforced on December 1st, 2013, on the behavior of algorithmic market participants involved in HFT activities. However, we believe that imposing a minimum time between the submission and cancellation of a limit order would represent a real solution to the issues raised in this chapter, which unfortunately is not part of the current regulations. Consequently, it is not impossible for technological advances to lead to even shorter duration orders providing less liquidity.

5.5 Algorithmic and High-frequency Trading Regulations in Germany

As claimed before, during our data period going from February 1st to April 30th, 2013, the regulations were still very light regarding algorithmic and high-frequency trading activities, but a certain transition was clearly perceptible on the horizon. Indeed, on February 28th, 2013, the *German Bundestag* approved the *Act for the Prevention of Risks and the Abuse of High-Frequency Trading*, which is now known as the *German High Frequency Trading Act* (*German HFT act* hereafter). The main goal of this act was to address measures and procedures to regulate algorithmic trading and high frequency trading (HFT). In fact, following the May 6, 2010, US flash crash, several authorities have looked into ways to reduce the risks related to algorithmic and HFT activities. Germany was the first country to present and adopt a concrete law in this direction. It applied to any firm involved in high-frequency algorithmic trading techniques in this country, whether local or foreign. As reported in Xetra Circular 073/13 on July 19, 2013 : *a high-frequency* *algorithmic trading technique is characterized by infrastructures that intend to minimize latency, as well as system determination of order initiation, generating, routing or execution without human intervention for individual trades or orders and by high intraday message rates which constitute orders, quotes, or cancellation.* Given this definition and those that will follow, up to some extent, all firms using algorithms in their trading activity have been affected by these regulations. In the specific case of Xetra market participants, the measures required by this legislation have been implemented over the period going from May 15, 2013, to April 1, 2014. The new requirements has been communicated to these members through a series of circulars of which we will summarize the main points.

The first measure imposed by the German HFT act is the requirement for any firm involved in HFT activities to obtain a license from the German Federal Financial Supervisory Authority (BaFin). While already required for financial institutions and financial services companies, a BaFin license became also mandatory for indirect trading participants such as market participants' clients performing HFT (Xetra Circular 045/13). Consequently, on May 15, 2013, market participants who had not obtained such a license and had not taken steps to obtain one had to quit their HFT activities. For participants in the process of obtaining this license, transitional periods of 6 and 9 months were allocated, depending on whether they were based in Germany or elsewhere. During these periods respectively ending on November 14, 2013, and February 14, 2014, the concerned participants were allowed to continue their HFT activities in accordance with the new regulations. Haferkorn and Zimmermann (2014) have used these three dates in an events study regarding the impact of the German HFT act. Focusing on 26 of the DAX 30 index components, they have observed that while the number of executed transactions have remained similar, the number of orders submissions have decreased, mostly following May 15, 2013, and February 14, 2013. Since orders submissions are often related to HFT, they have related these decreases to previously unregulated firms who may have ceased or adapted their Germany stock markets trading activities rather than attempted to meet the BaFin license requirements. They have related the decrease observed after the May 15, 2013, to firms preferring to abandon HFT trading rather than take even a single step in an attempt to obtain this license. On the other hand, they linked the decrease following

February 14, 2013, to foreign firms that may have chosen to abandon the registration process after the 9-month transitional period that was allocated to them.

Regarding the market participants legally authorized to perform HFT activities after the adoption of the German HFT act, they are now subject to several constraints in terms of organizational obligations, capital requirements, reporting, disclosure, risk management and system reliability (Haferkorn and Zimmermann (2014)). When it comes to actual trading activities, three main measures have been implemented. First, it has become required for the market participants to identify any order produced by an algorithm (Algo-Flagging). As reported in Xetra Circular 099/13, with regard to this measure : *The Trading Participants must flag orders and binding quotes generated by them through algorithmic trading and identify the trading algorithm used*. In slightly more detail : *orders or quotes generated, modified, or deleted via algorithms and the identification of the trading algorithms must be flagged*. Therefore, each algorithm now has to be clearly identified by *a single unambiguous key* that remains unchanged during its lifetime. This key then has to be used in order to identify the actions performed by the algorithm. In the specific case of Xetra, after a transitional period during which systems had to be modified to be able to meet these requirements, market participants were obligated to apply these rules on April 1, 2014. The second measure was the imposition of Order to Trade Ratios with the goal of ensuring that the Market Participants guarantee an appropriate ratio of their orders entries, modifications and cancellations to the transactions actually executed. These ratios are calculated on a monthly basis and the reset takes place at the beginning of the month. Similarly, the third measure was the imposition of fees for Excessive system usage, with the goal of discouraging disproportionally high number of order entries, modifications, and deletions. The application or not of such charges is determined on the basis of daily activities. For Xetra, both these measures have been totally enforced from December 1, 2013.

As far as we are concerned, it is these last two measures that are likely to have had the greatest impact on the nature of the activities observed on Xetra. Indeed, since they involve actual economic consequences in the form of penalties, we consider that they had an important potential of modifying the behavior of the market participants previously engaged in almost unlimited HFT activities. However, we must remain cautious about the real scope of these measures. Indeed, imposing some limits on HFT, the German HFT act does not explicitly prohibit any activity. Our way of seeing things it that, because of the Order-to-trade ratios, the concerned firms end up with a monthly *HFT budget* to be respected in terms of orders submissions, modifications, and cancellations. Insofar as this budget depends on the number of transactions actually executed, we can assume that all these firms are not equal in absolute terms. Thus, it is very likely that a major financial institution producing a high volume of transactions that could be generated by it brokerage and portfolio management activities, could end up with an important allowed number of orders submissions, modifications, and cancellations. This could prove to be an advantage if it also participates in HFT activities, which would be less limited than those of a firm with more restricted operations. Since these *HFT budgets* depend on the number of transactions executed, it is also possible to assume that they are potentially not constant from month to month. Indeed, we consider highly possible that, given the fact that actual orders execution may be related to portfolio rebalancing and similar activities, the number of these operations and consequently, the resulting order-to-trade ratios, may vary from a month to another. In this context, forecasting and managing these *HFT budgets* has potentially become an important task in order to maximize HFT activities without exceeding the monthly limits. It is even relevant to ask if certain firms decrease or increase their HFT activities at the end of the month in order to respect these ratios or, on the contrary, to approach them as much as possible.

We consider likely that the date of December 1, 2013, could have actually been a turning point in this regard on Xetra. In order to determine whether any structural change has taken place, it would be very relevant to repeat the analyzes produced in this chapter on Xetra data which come from a period subsequent to this date. However, since we do not have such data, we must leave this work for future consideration.

5.6 Summary

In this chapter, we have used the data obtained by implementing our Chapter 3 orders tracking methodology to perform various analysis. We have mainly focused on the successfully tracked orders whose life cycle appears to end with a cancellation for the some of the DAX, MDAX and SDAX Xetra indexes components over the time period going from February 1 to April 30, 2013.

In a first general analysis in which we have considered the limit orders as individual liquidity vehicles, we find that for the DAX, MDAX and SDAX stocks of interest :

- 30.6%, 33.6% and 38.1% of all orders are canceled without having moved a single position with regard to their order book side best price. This led us to question the intentions behind the submission and cancelation of limit orders. Indeed, by definition, patient investors would submit limit orders on the book with the hope of a favorable price movement that could lead to their execution. However, from our point of view, cancelling a limit order that have not moved a single position after its cancellation is somewhat inconsistent with this prior.

- 31%, 35% and 51% of all limit orders are cancelled less than one second after their submission. Given the human being limitations on a sustainable basis, we identify these orders as part of algorithmic and high-frequency trading activities.

- 10.6%, 16.1% and 13.3% of orders are submitted and cancelled without real execution possibility. We have obtained the threshold probabilities by computing execution statistics based on price level of submission and executed orders time between submission and execution.

In the second part of this chapter, after having observed that 88.7%, 90.2% and 93.9% of the limit orders cancelled less than one second after their submission appear closely followed by the submission and cancellation of other orders sharing similar characteristics, we have put the emphasis on short duration orders sequences. Focusing on the SDAX stocks for which we consider the orders data as highly reliable, we actually find that 11.5% of all limit orders are part of sequences involving the submission and cancellation of 100 limit orders and more. We relate these sequences to potential quote-stuffing operations. It is highly interesting to note that the orders involved in sequences that consist in 10 orders and more account for 38.2% of all the limit orders observed for these stocks.

Given these facts, we have developed a frame of reference allowing to classify and potentially identify the algorithms that we consider as involved in quote-stuffing activities. By looking at the orders sequences at a microscopic level for the SDAX index components, we have revealed a world that exceeded our prevailing expectations. In fact, we have observed that in multiple cases, orders are submitted and canceled at a rapid rate that can only be orchestrated in an algorithmic context. We have been able to establish that it is very possible that certain algorithms are active on different stocks and that depending on their version or their parametrization, they may produce different results that share similar characteristics.

Finally, in a last analysis we have focused on the liquidity provided by our three orders duration categories with regard to the generated level of activity. Still focusing on the orders that we consider as algorithmically generated, we find that whether they are submitted in sequence or not, these orders globally provide 0.4%, 0.07% and 0.21% of all market liquidity for the DAX, MDAX and SDAX stocks analyzed. In counterpart, they contribute to 33.5%, 37.0% and 51.8% of all orders that we have successfully followed from submission to cancellation on depth levels 1 to 10. In absolute terms, we have shown that, as an example, for our three-month data period, the 293 913 orders with a duration smaller than 10 milliseconds observed for the SDAX IVG stock have provided as much liquidity as an average value order for a 7 minutes and 3 seconds period. From our point of view, this information tells us that a significant part of the liquidity observed in the Xetra stock markets for the time period of interest is simply virtual, particularly regarding the less liquid SDAX index components. Thus, we consider important that we, or other researchers, revisit these analyzes over a Xetra data period after 2013, as the German regulations have changed and should be more restrictive with regard to high frequency trading activities.

Figure 5.1 Passive orders submissions by price level

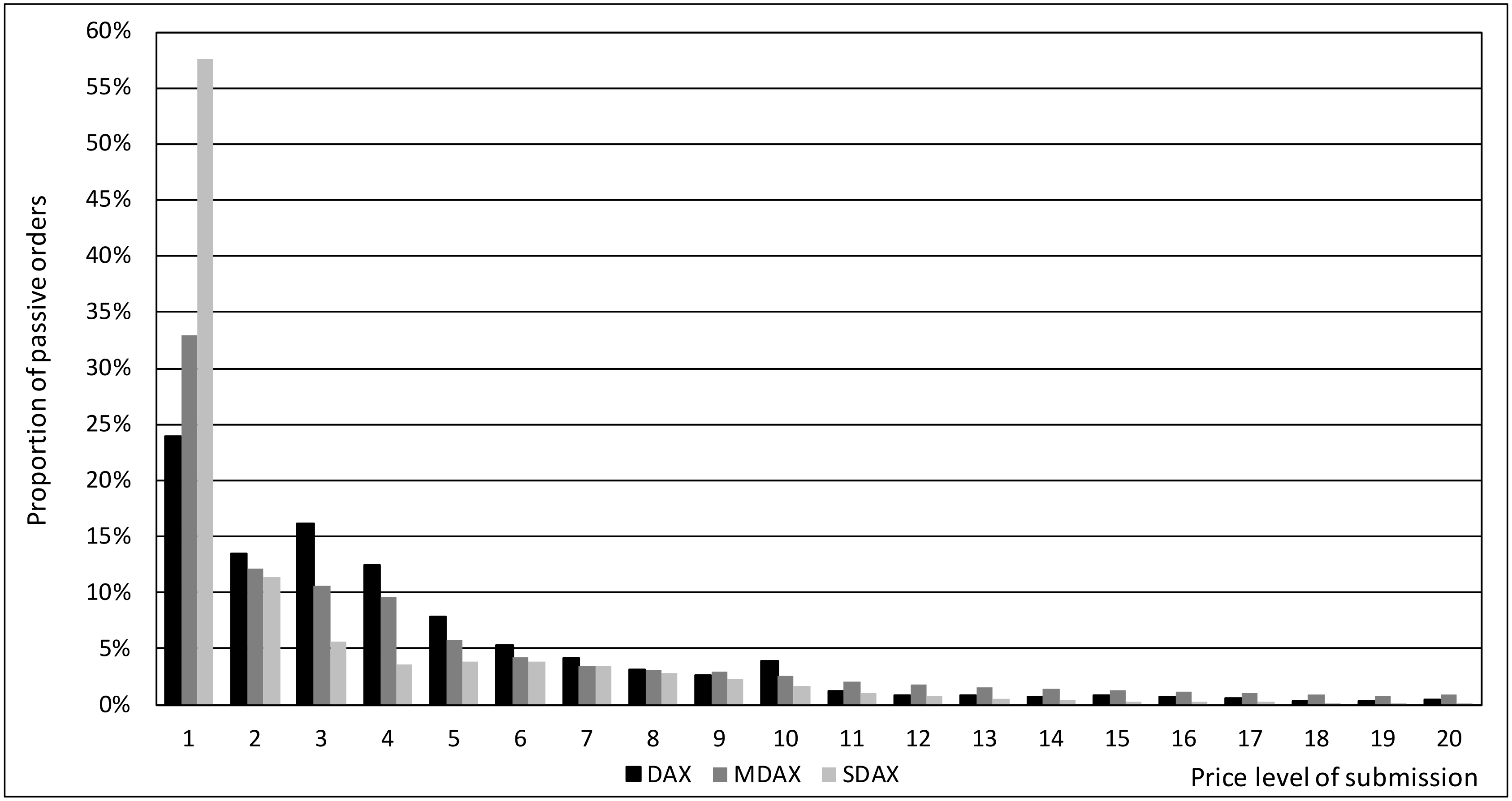


Figure 5.2 Passive orders cancellations by price level

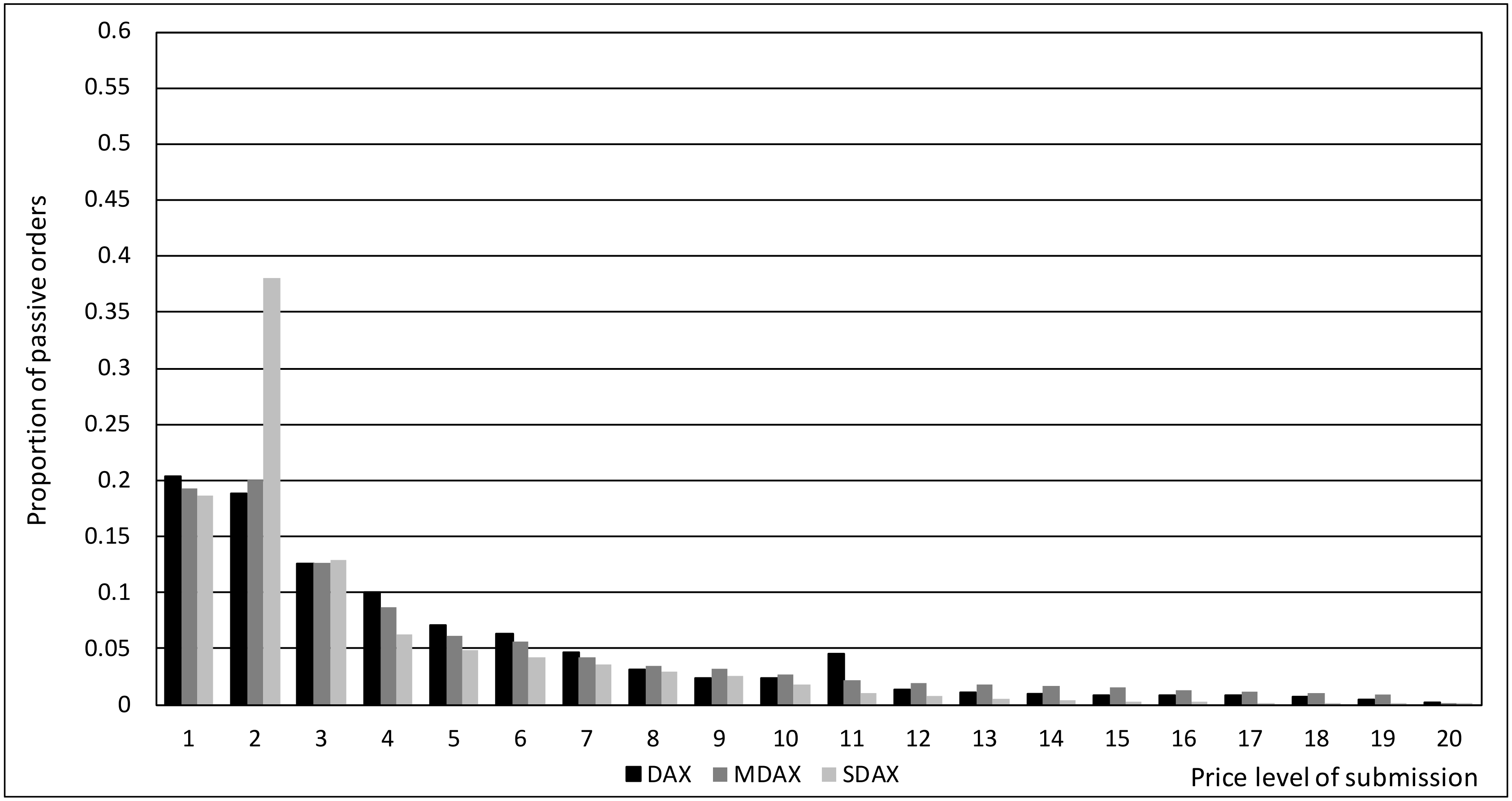
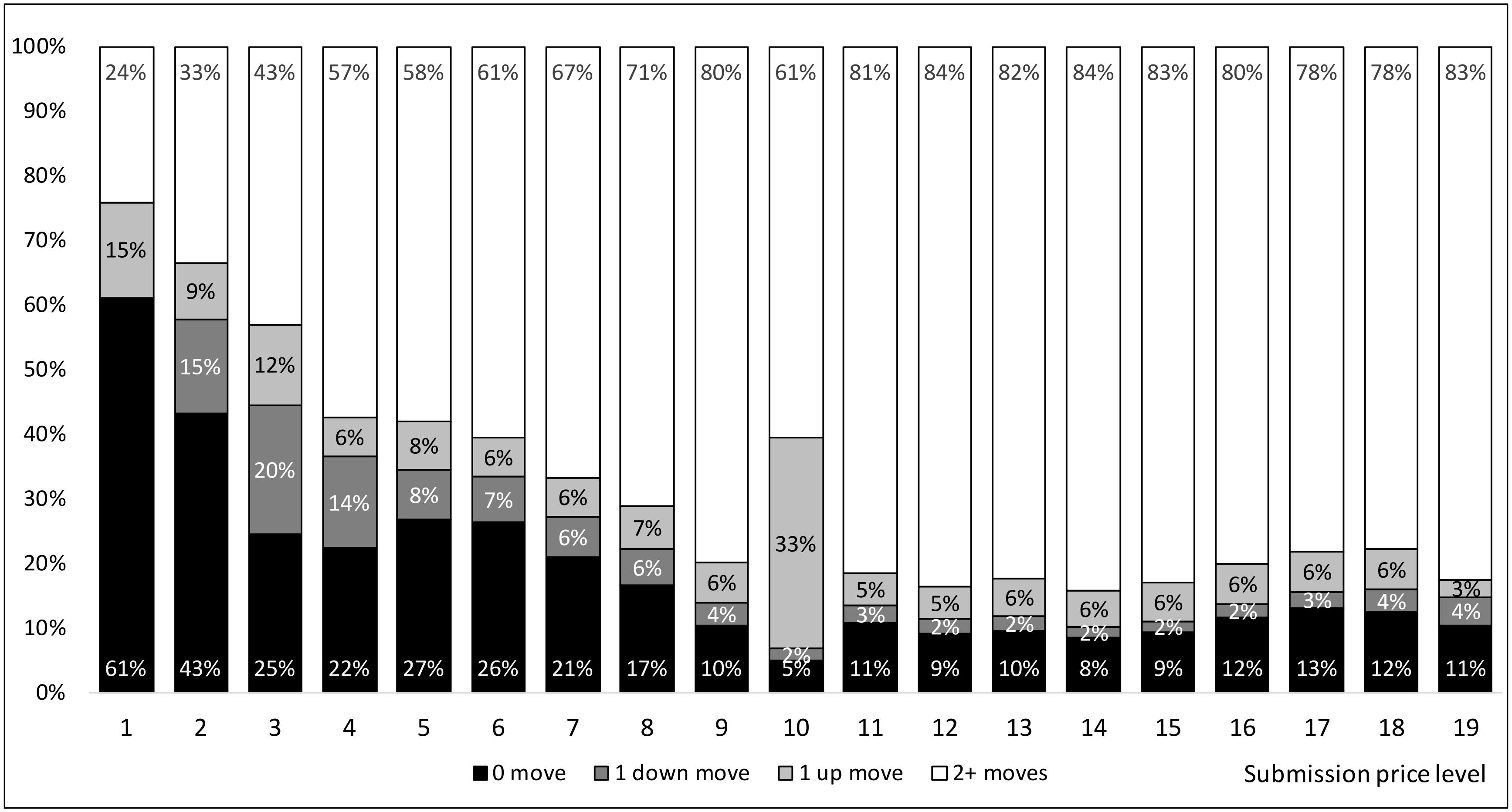
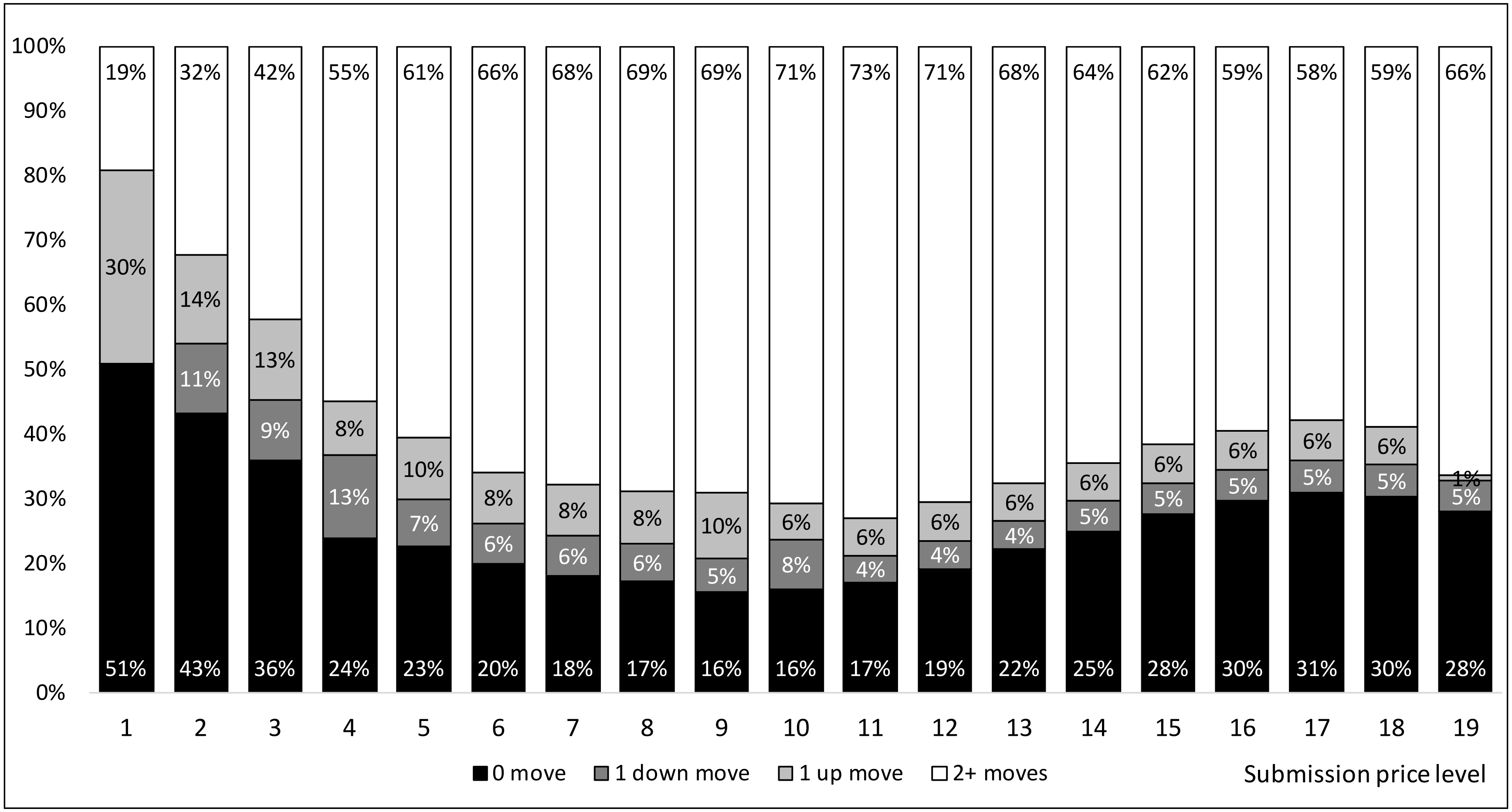


Figure 5.3 Position changes before cancellation

Panel I: ***DAX15*** **components**



Panel II: MDAX index components



Panel III: SDAX index components

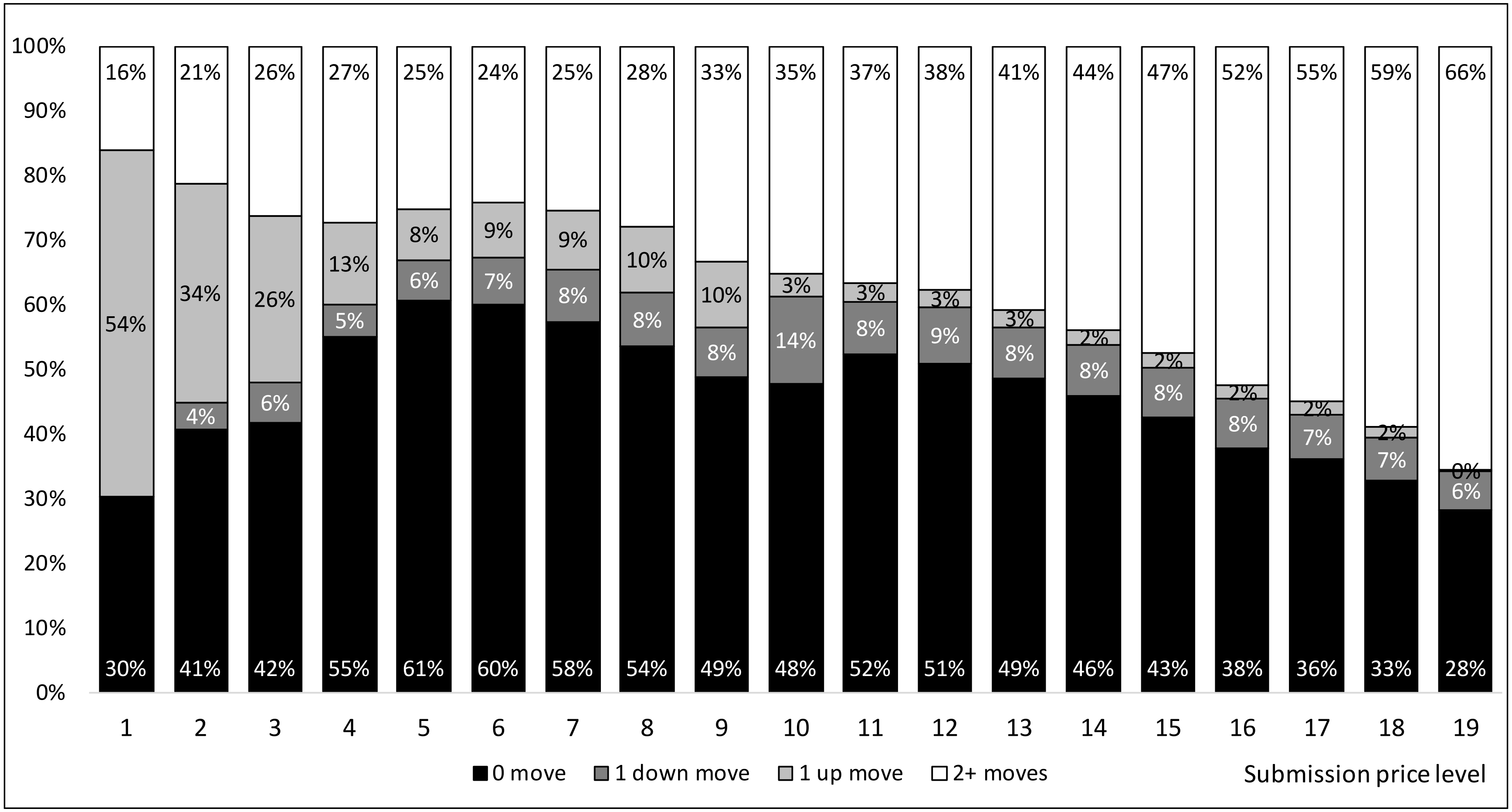


Figure 5.4 Cancelled orders duration CDF

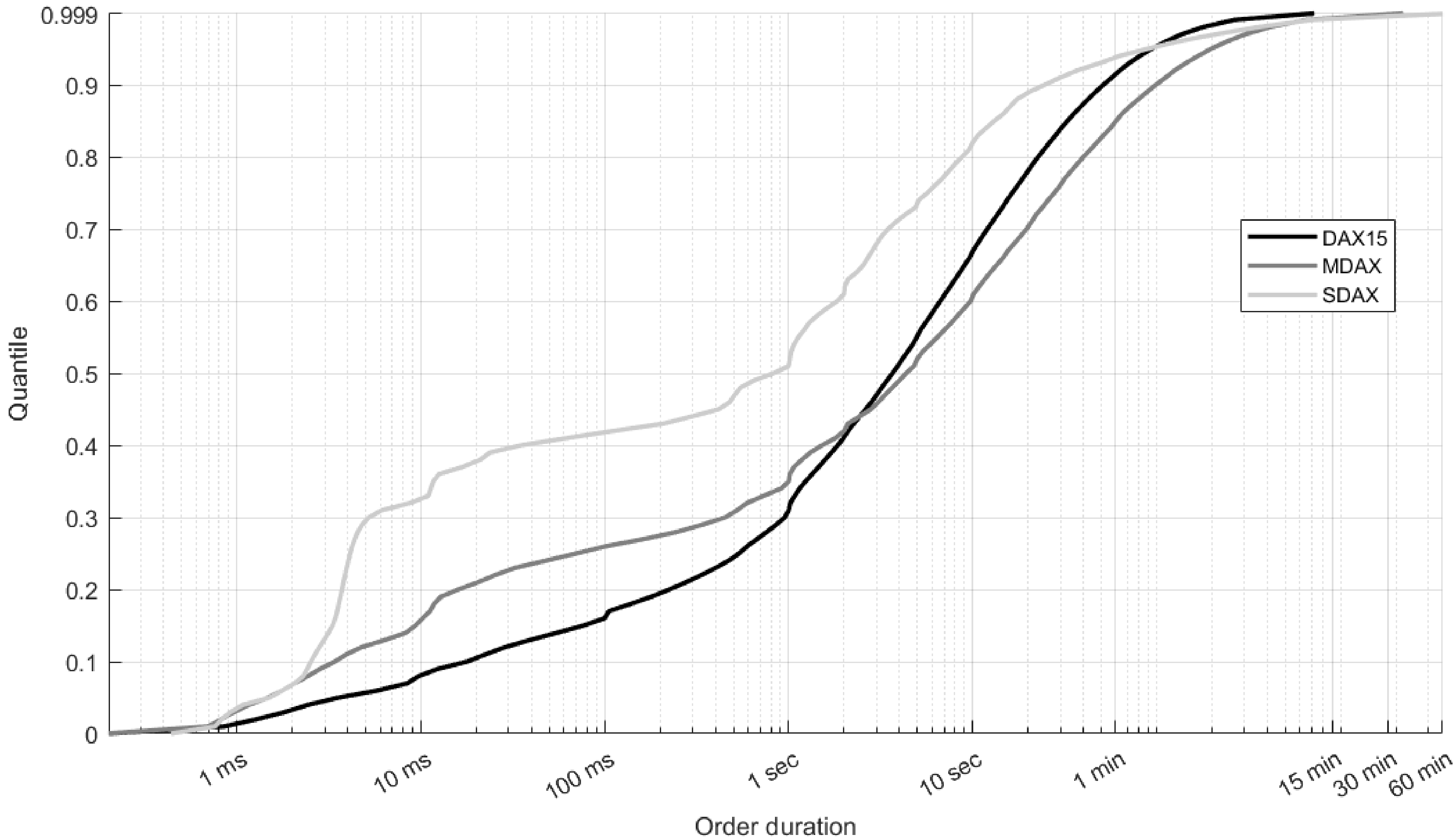


Figure 5.5 Cancelled orders duration by positions changes CDF

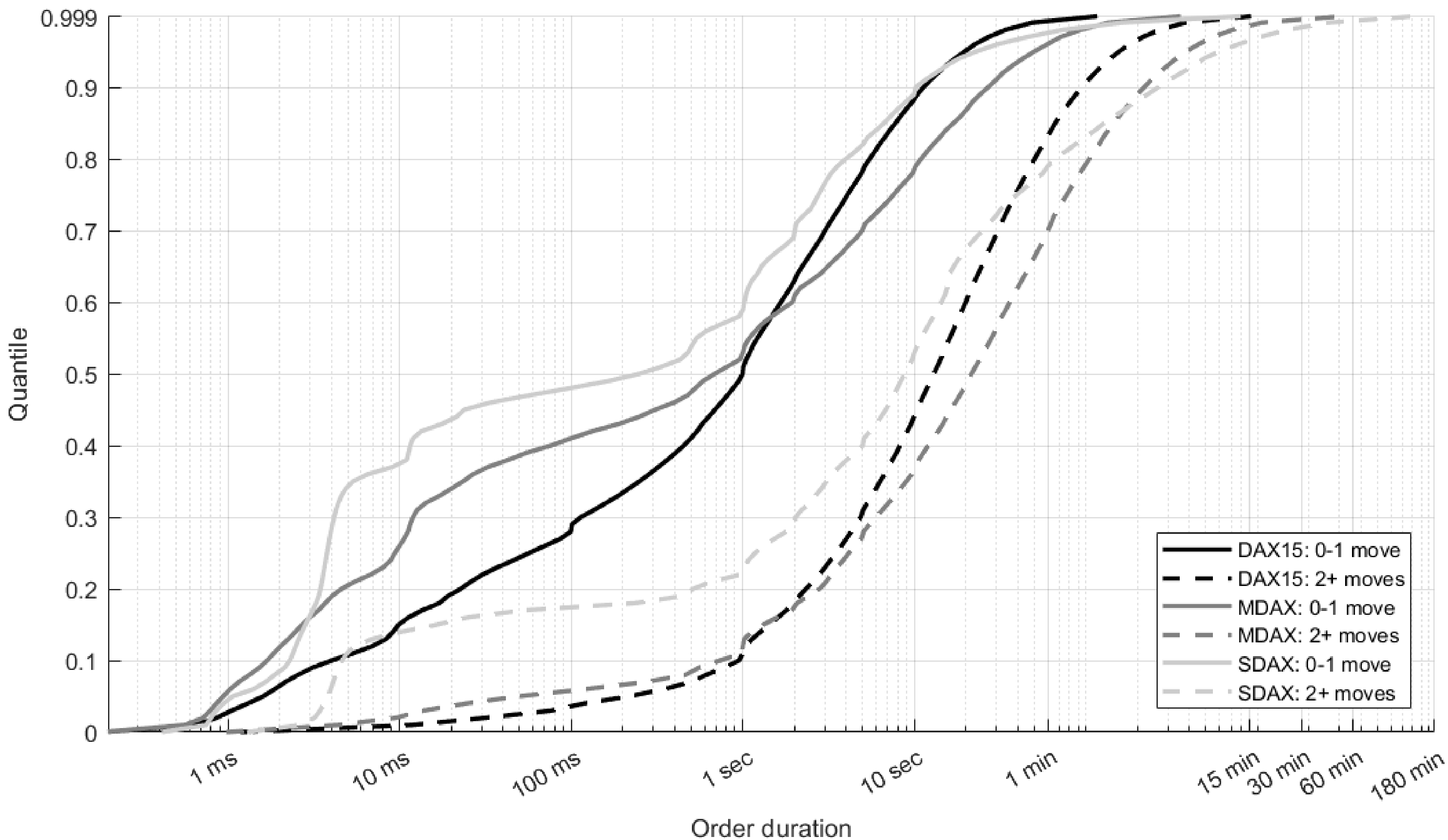
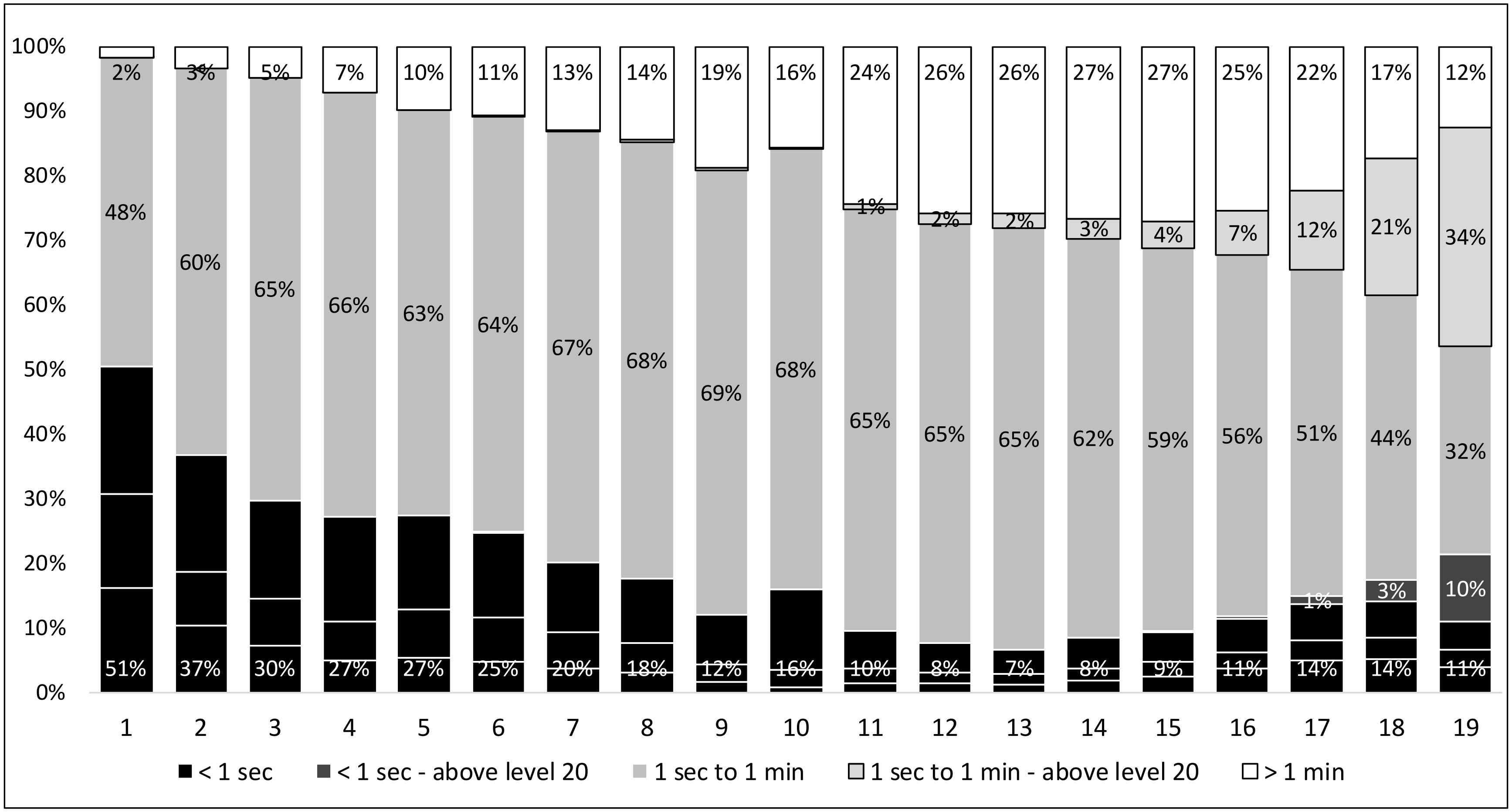
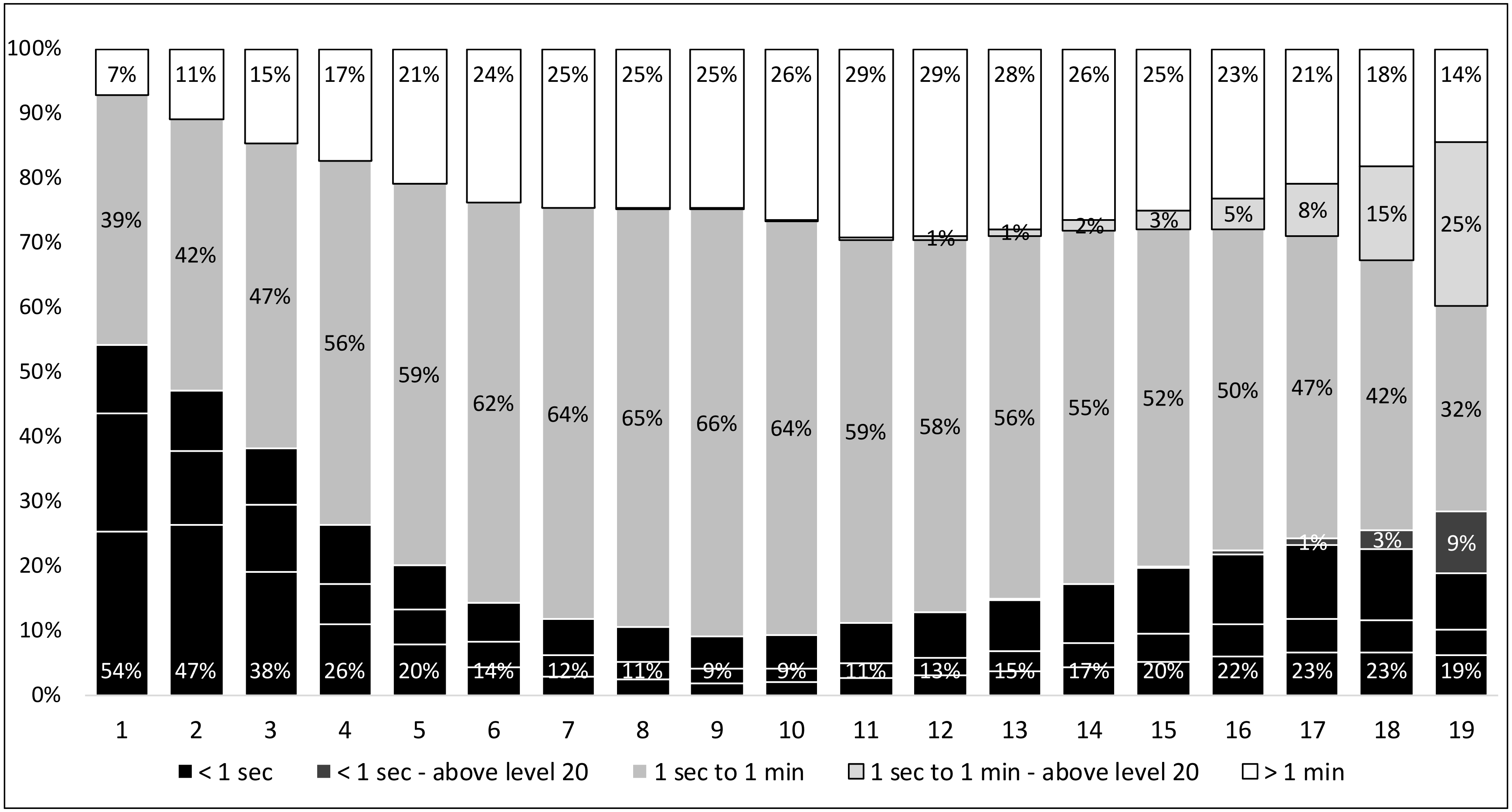


Figure 5.6 Duration before cancellation by submission price level

Panel I: *DAX15* components



Panel II: MDAX index components



Panel III: SDAX index components

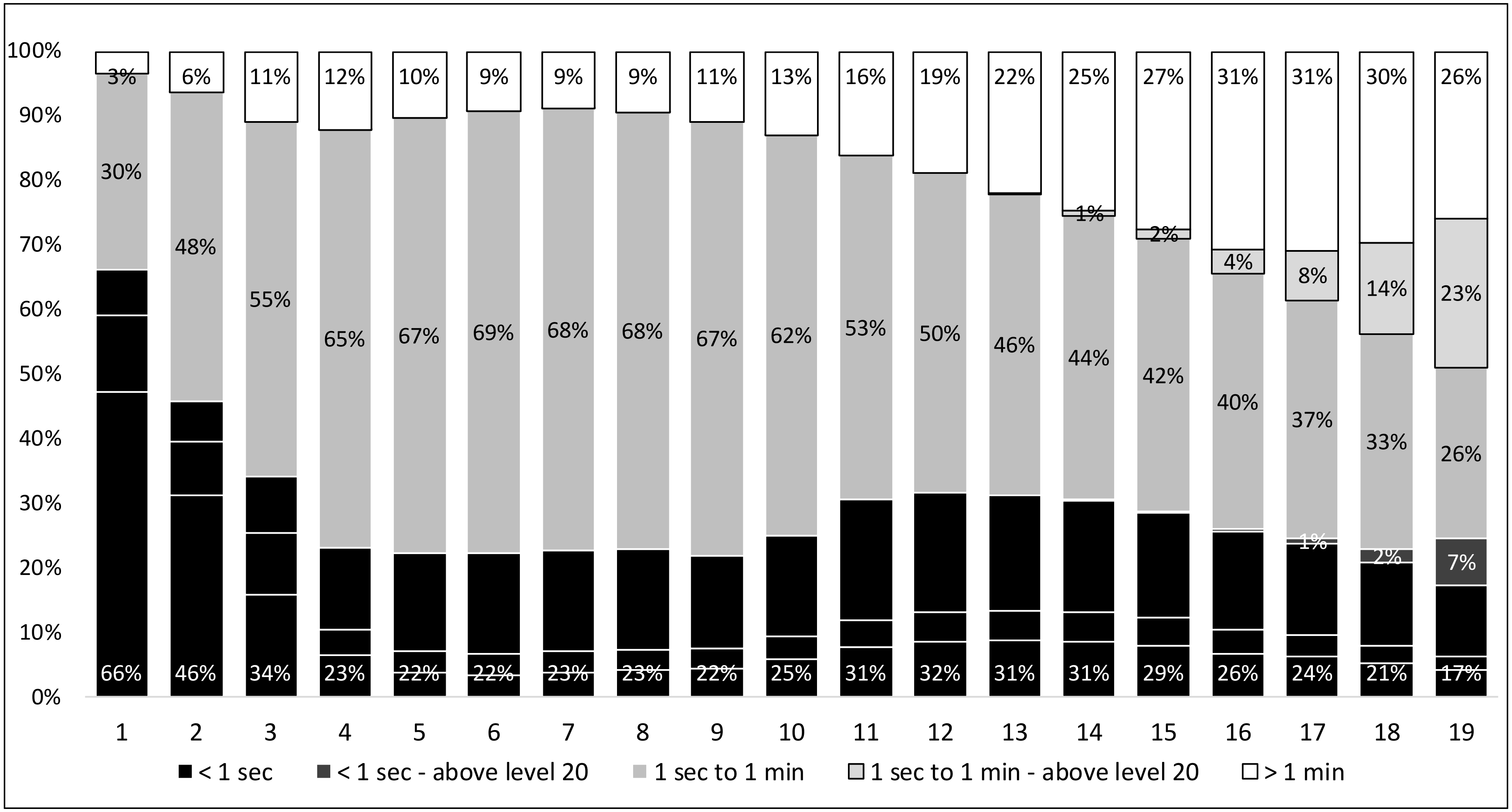


Figure 5.7 Totally executed orders duration CDF

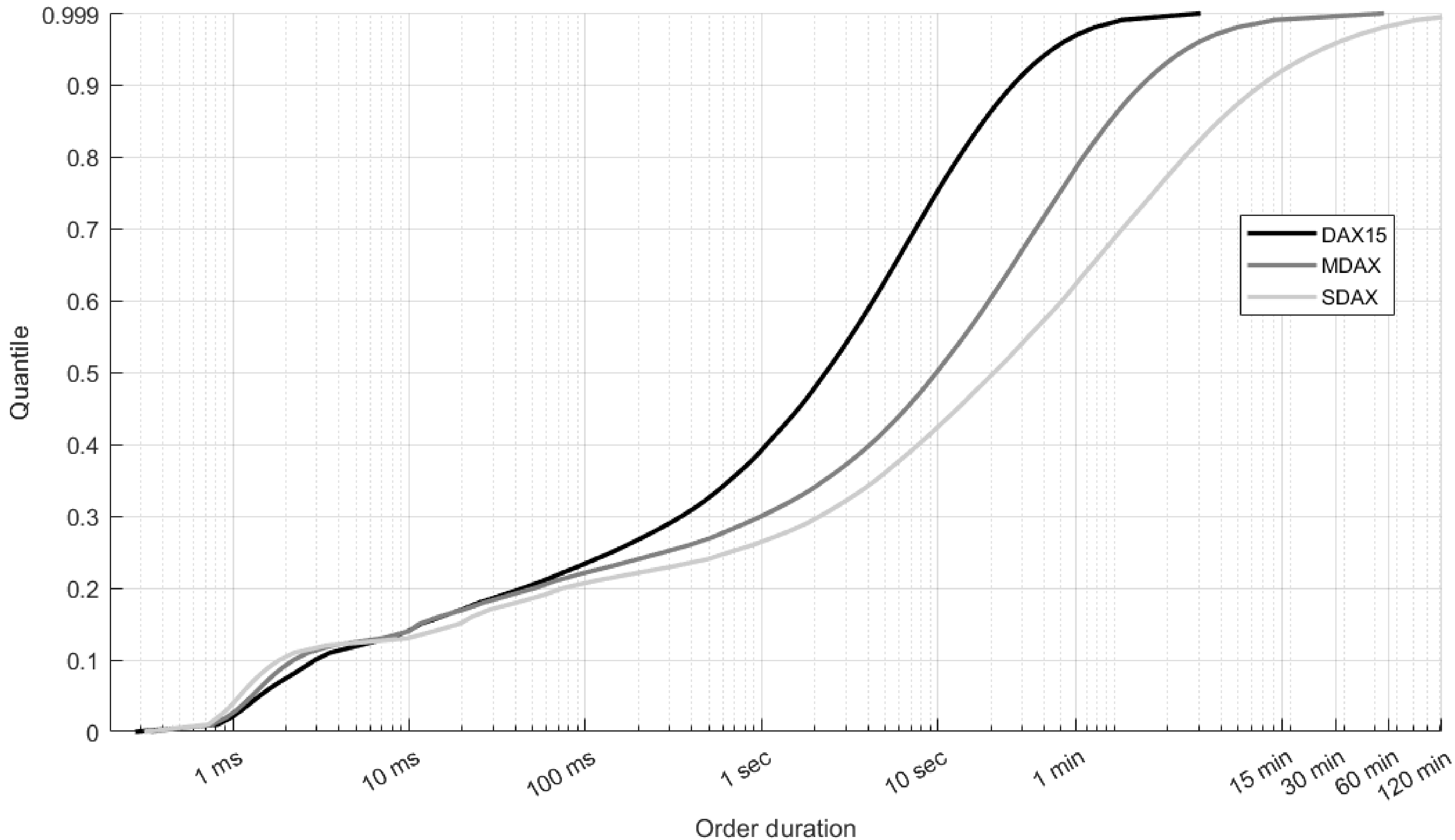
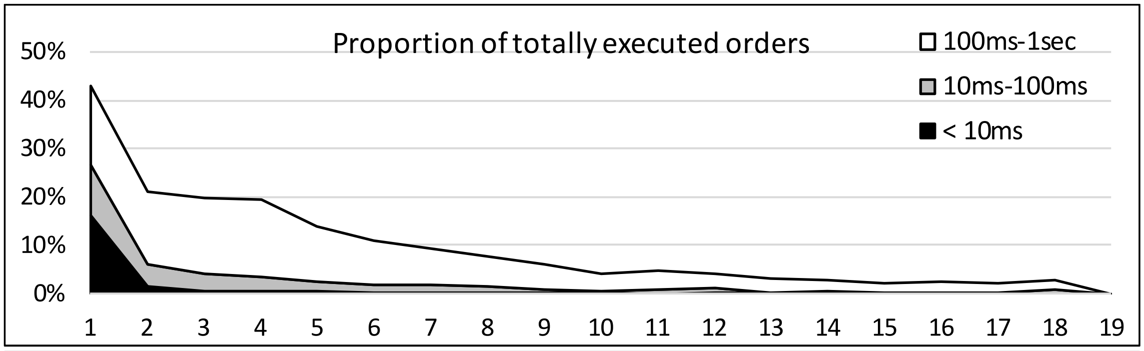
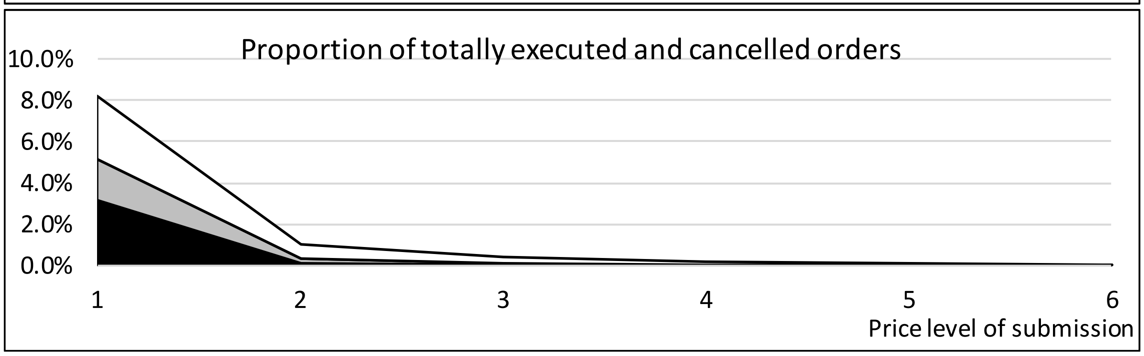


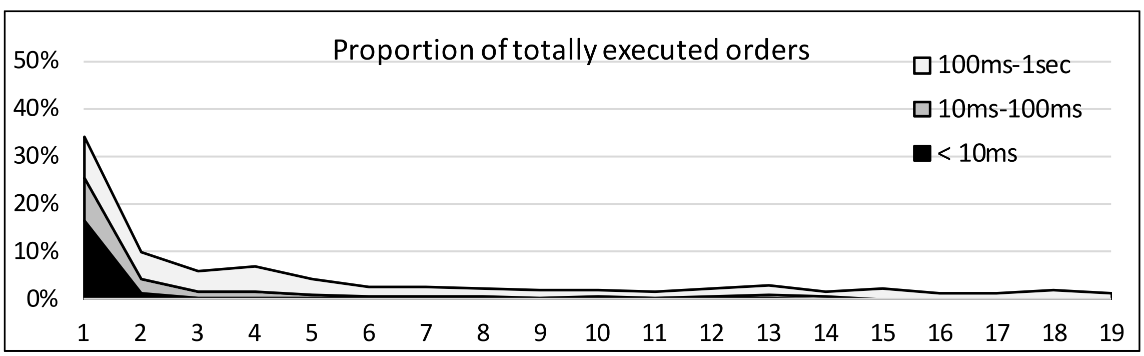
Figure 5.8 Totally executed orders in less than 1 second by price level of submission

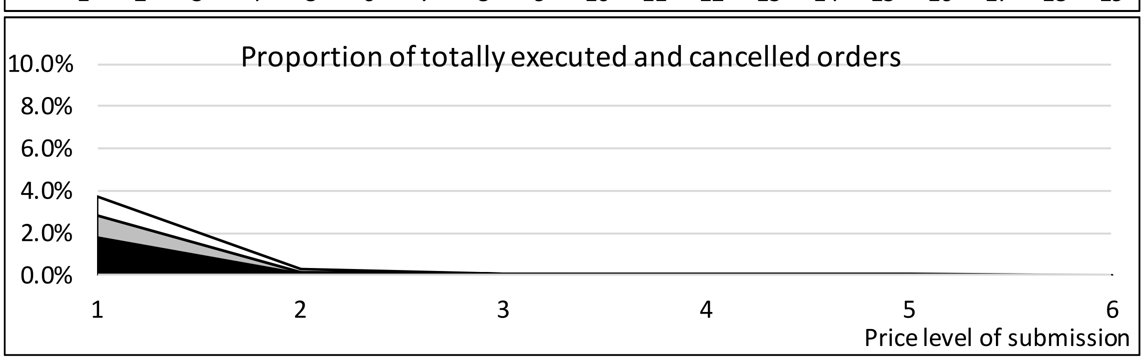
Panel I: *DAX15* components



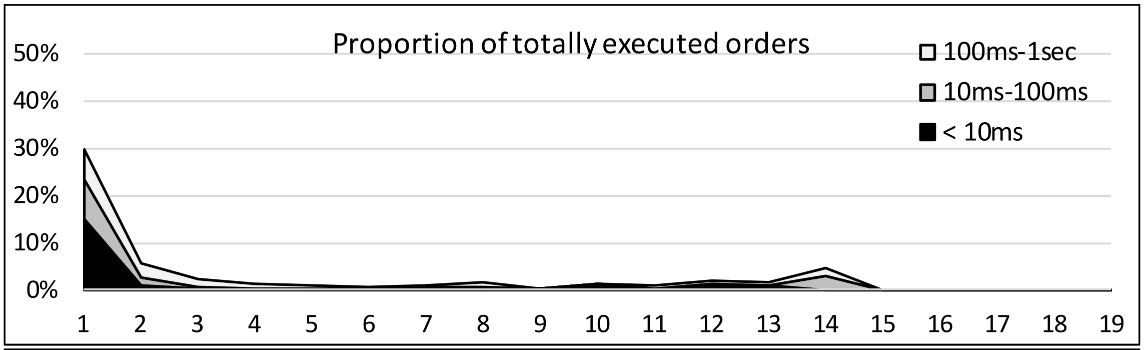


Panel II: MDAX index components





Panel III: SDAX index components



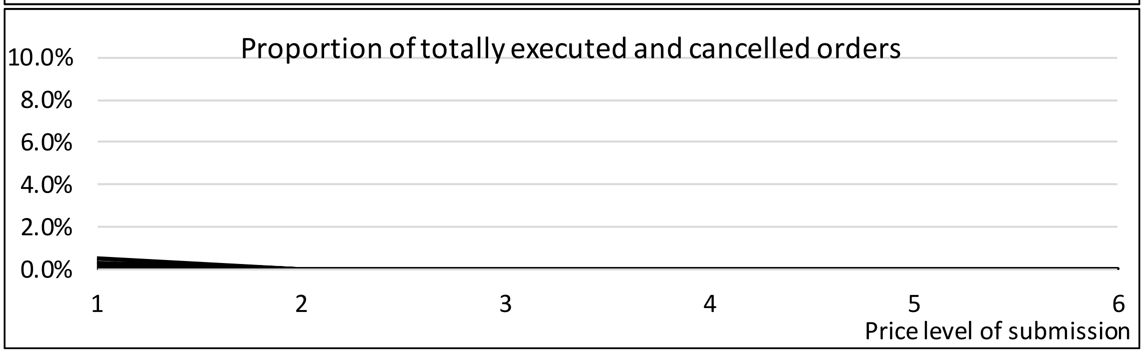


Figure 5.9 Interval between short duration orders submission CDF

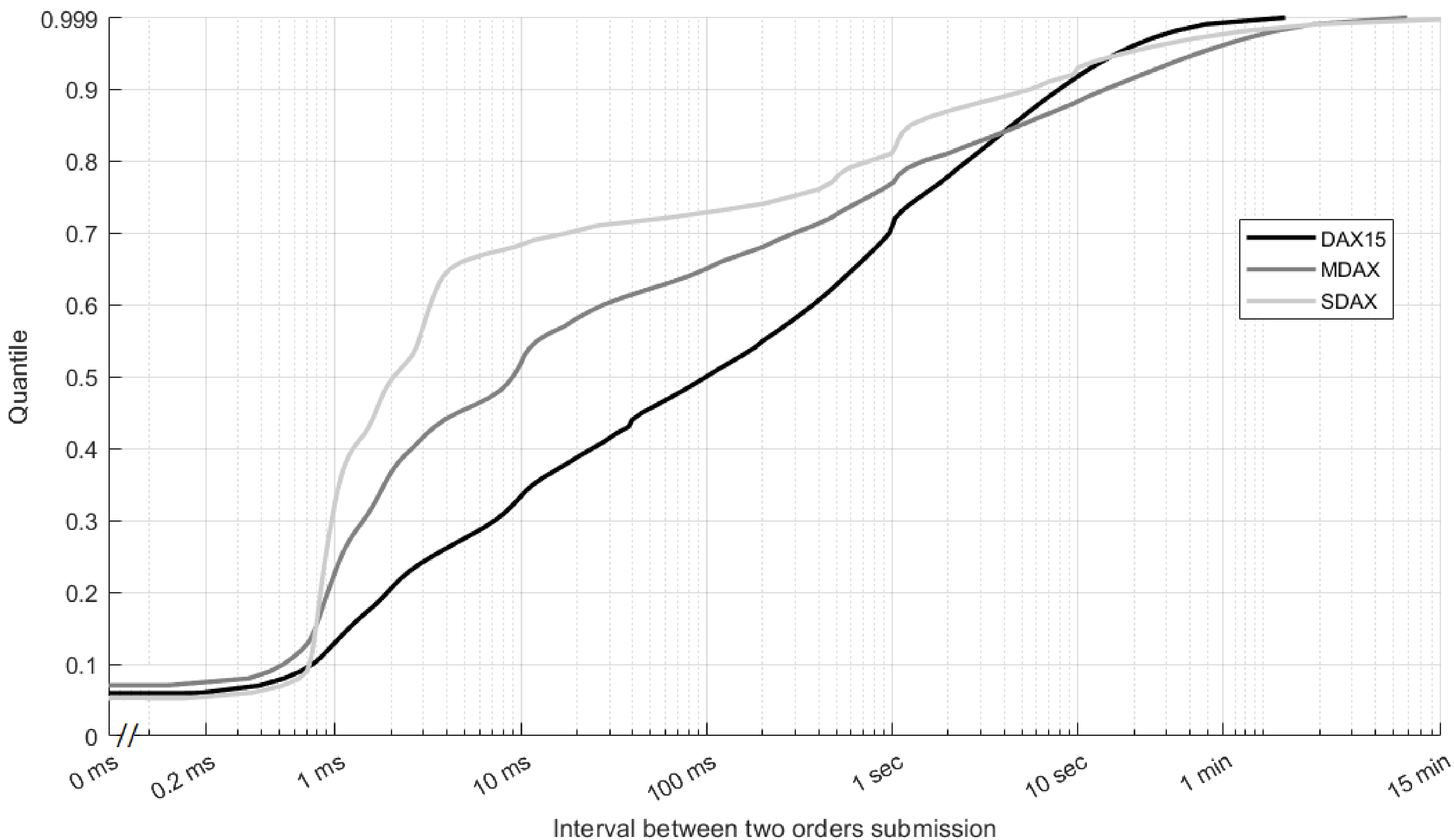
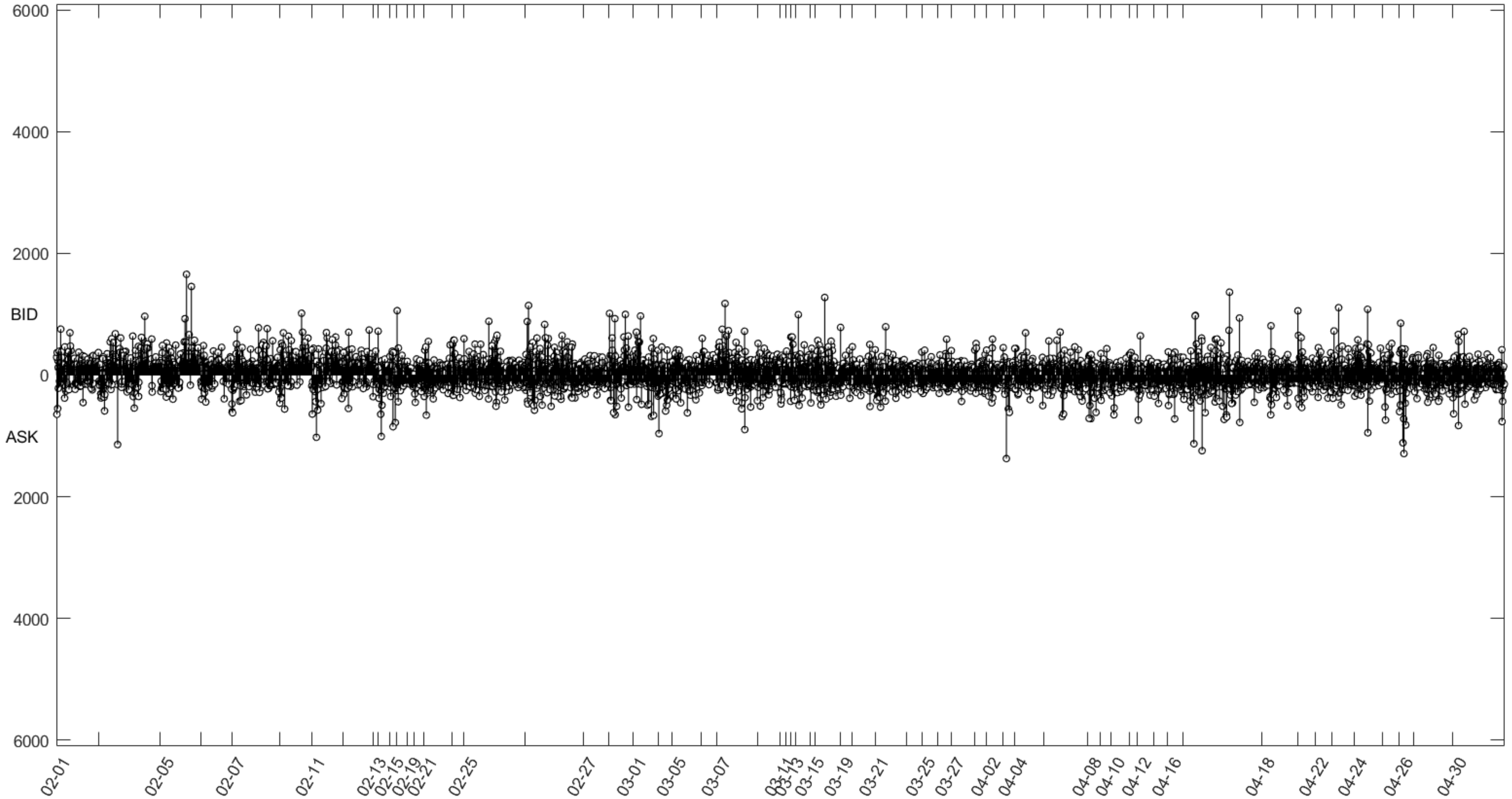
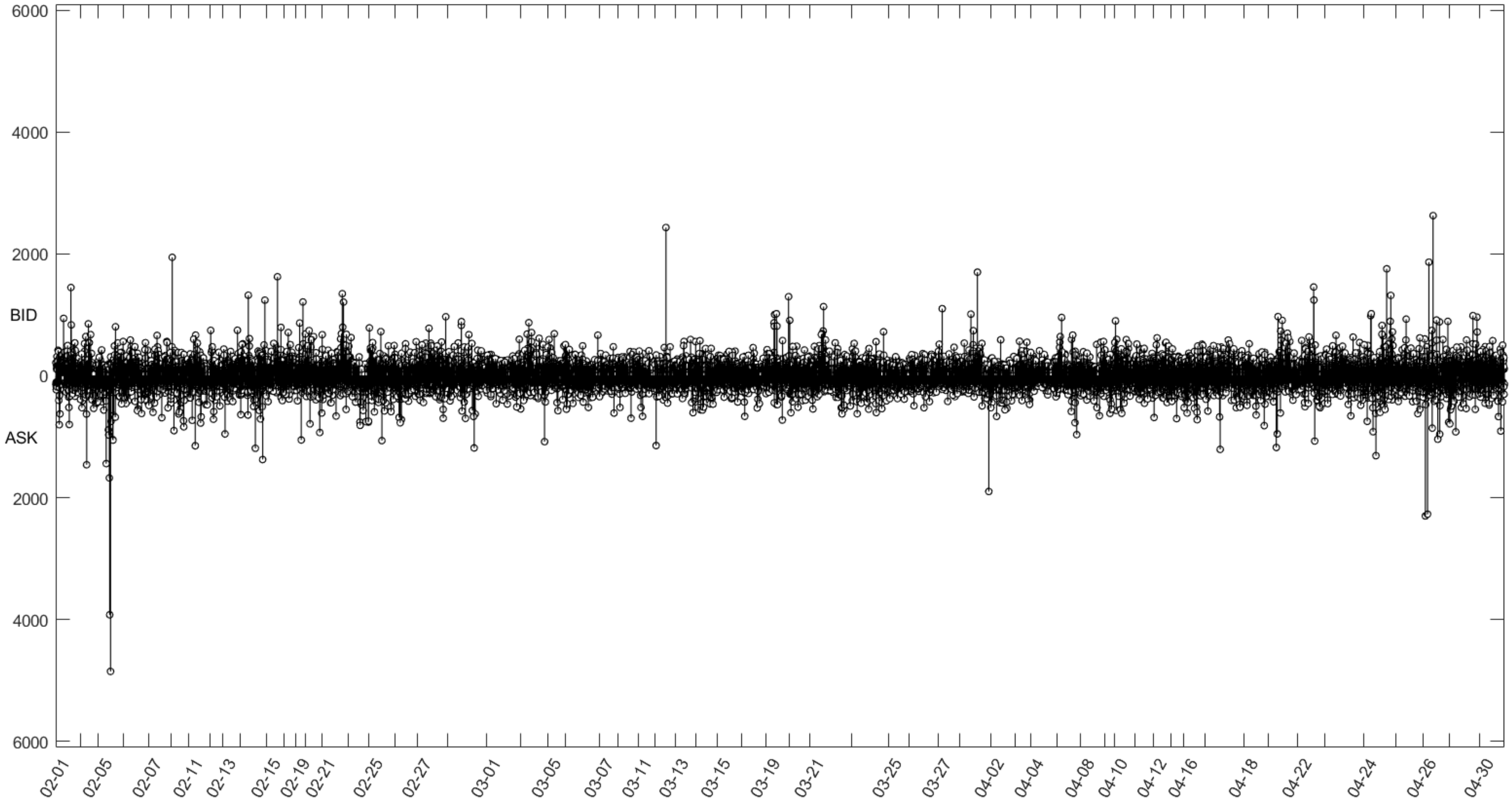


Figure 5.10 100 short duration orders and more sequences in time

Panel I: *DAX15* components



Panel II: MDAX components



Panel III: SDAX components

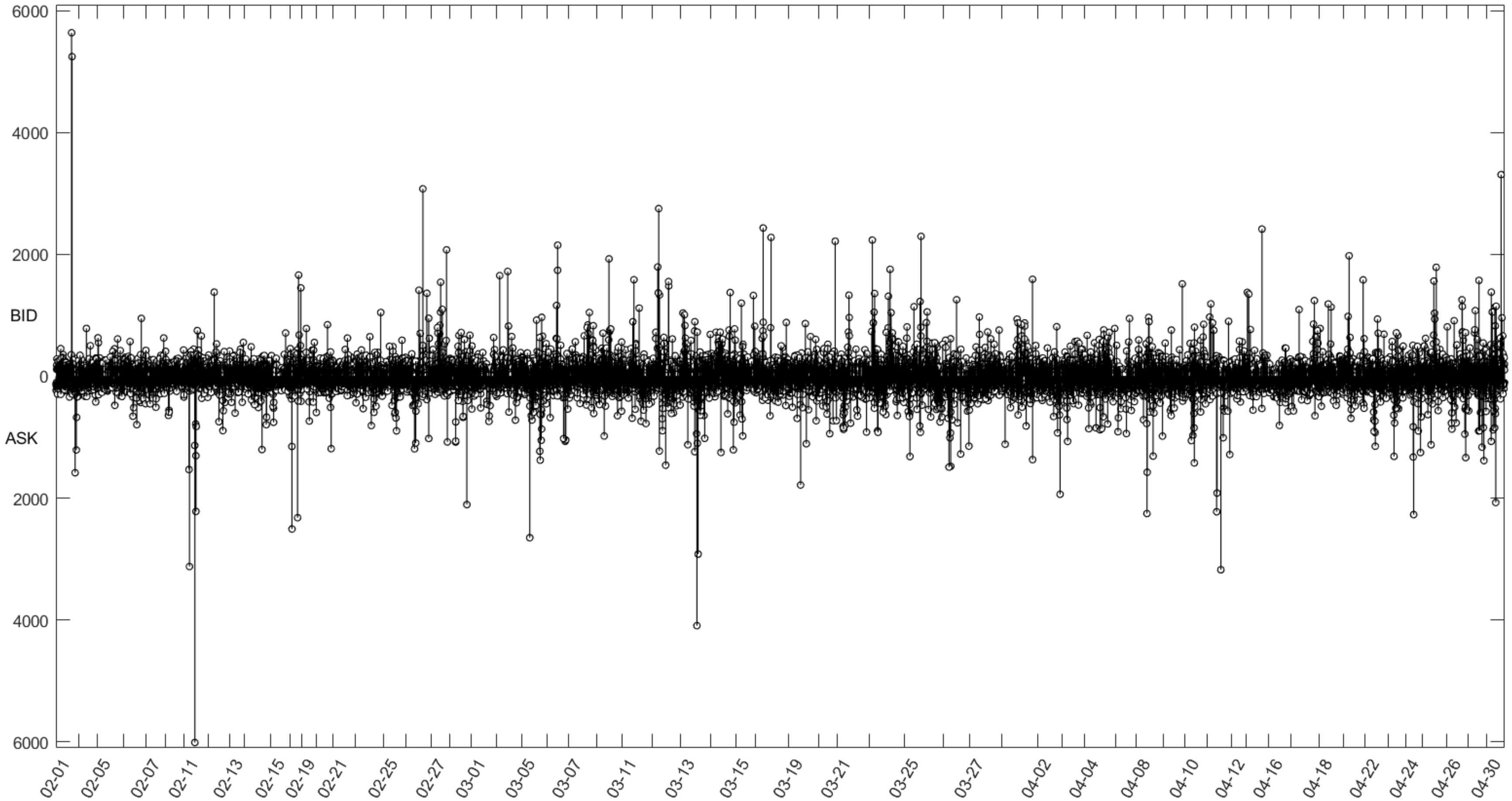
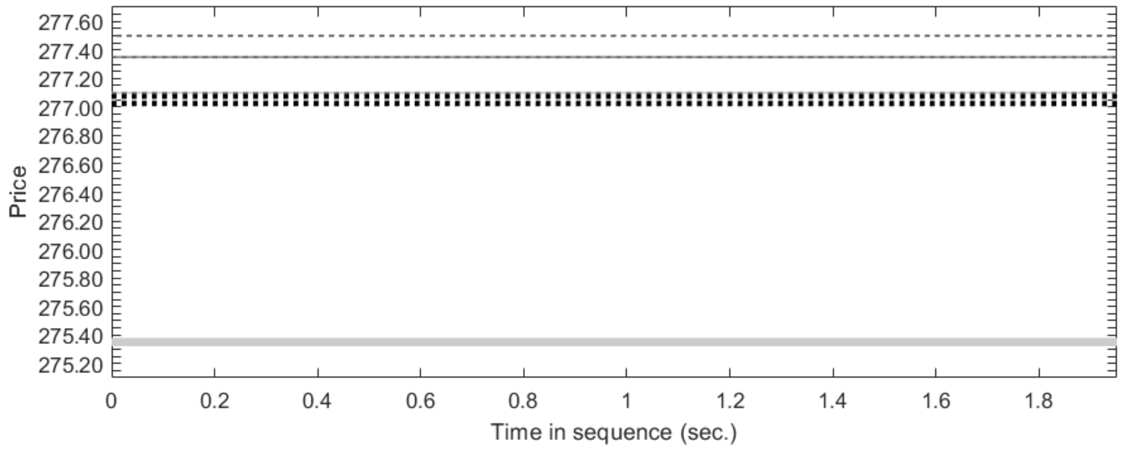


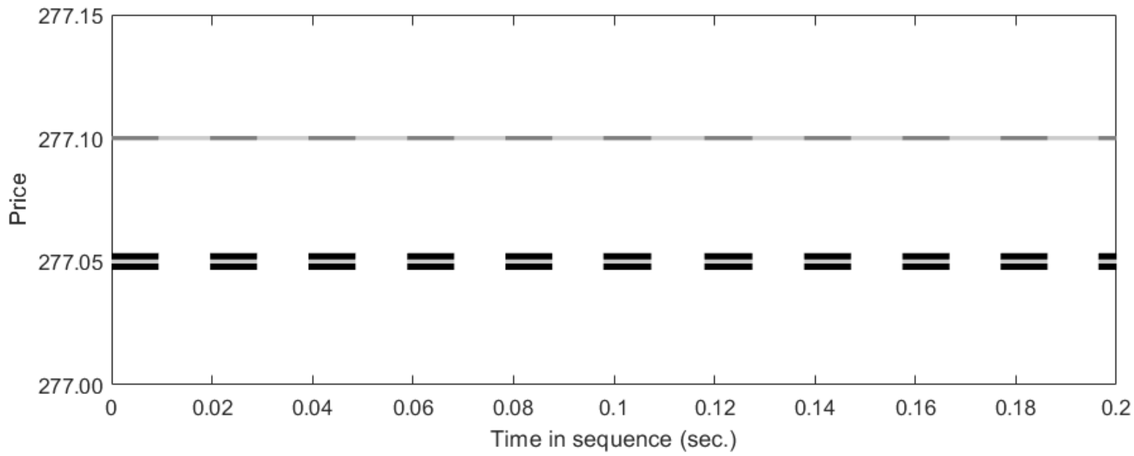
Figure 5.11 Orders Sequence general example

KWS stock on 2013-04-26 (15:05:01) – 100 ask orders

Panel A: Complete sequence – 1.95 seconds



Panel B: 0 to 200 milliseconds in sequence



Panel C: 0 to 40 milliseconds in sequence

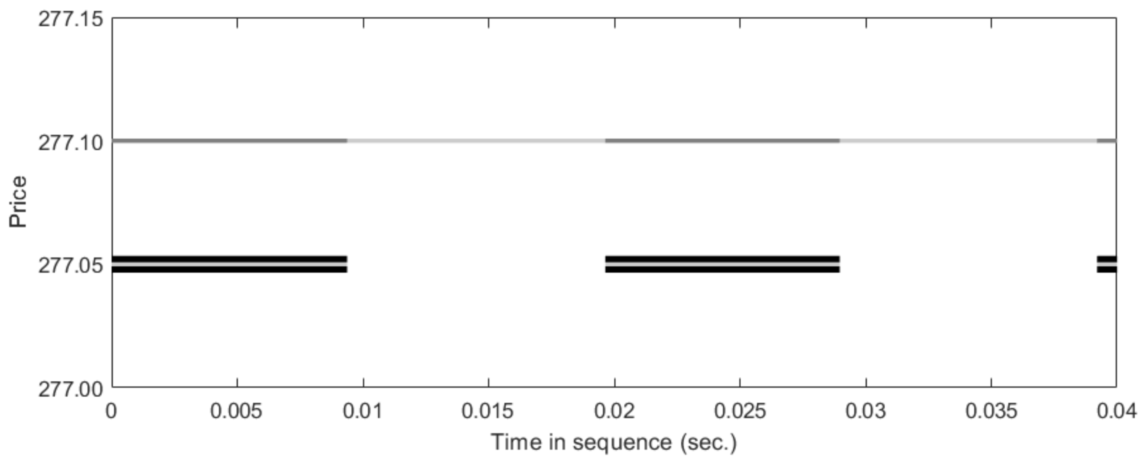
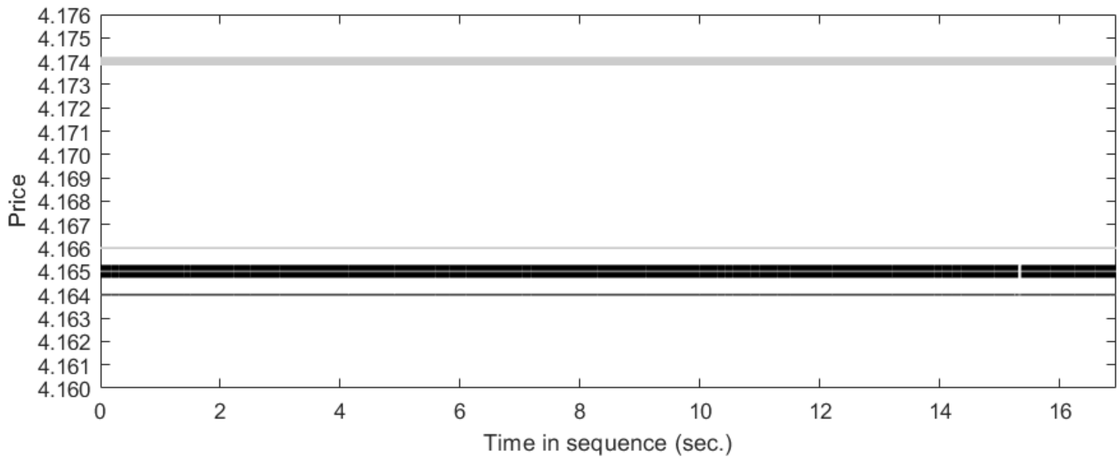


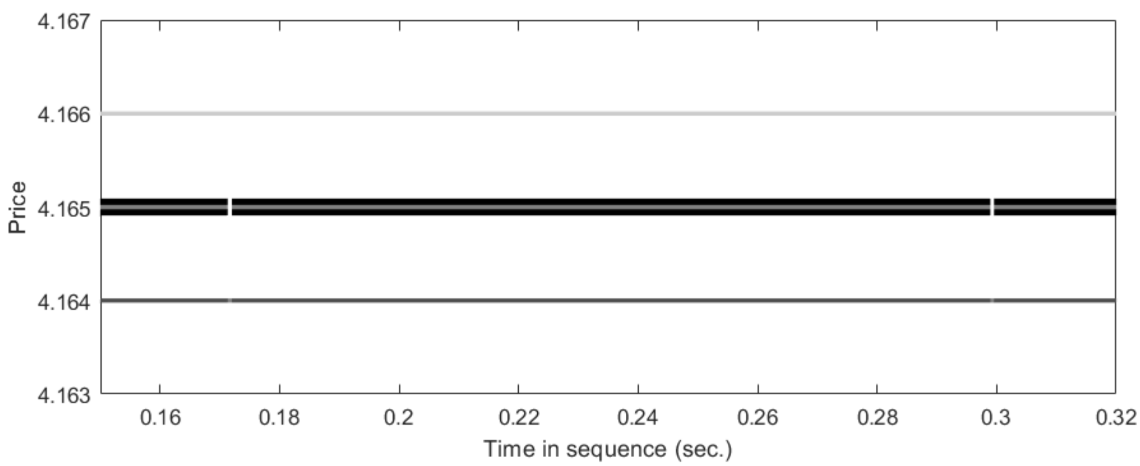
Figure 5.12 Algorithmic profiles examples

Example 1.2 – BAF stock on 2013-02-22 (12:25:49) – 122 bid orders

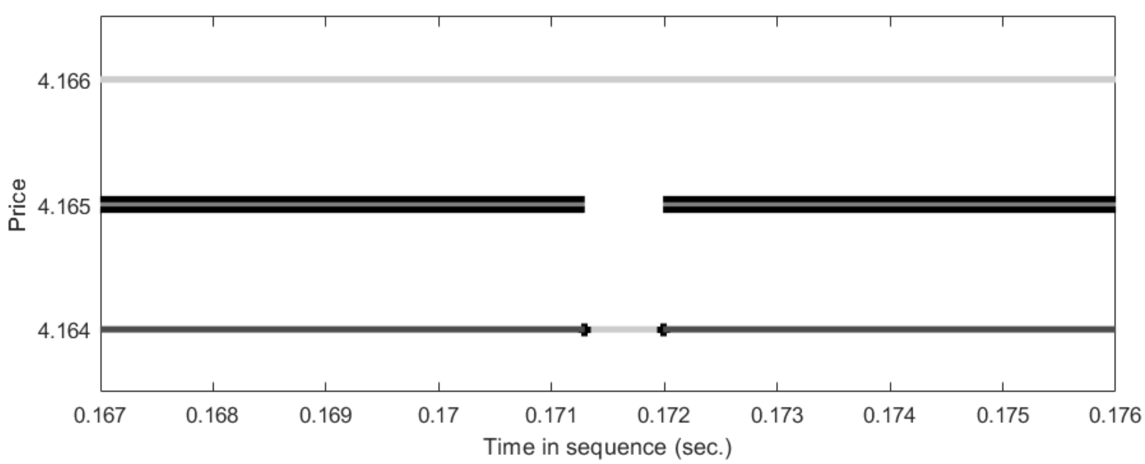
Panel A : Complete sequence – 16.94 seconds



Panel B : 150 to 330 milliseconds in sequence

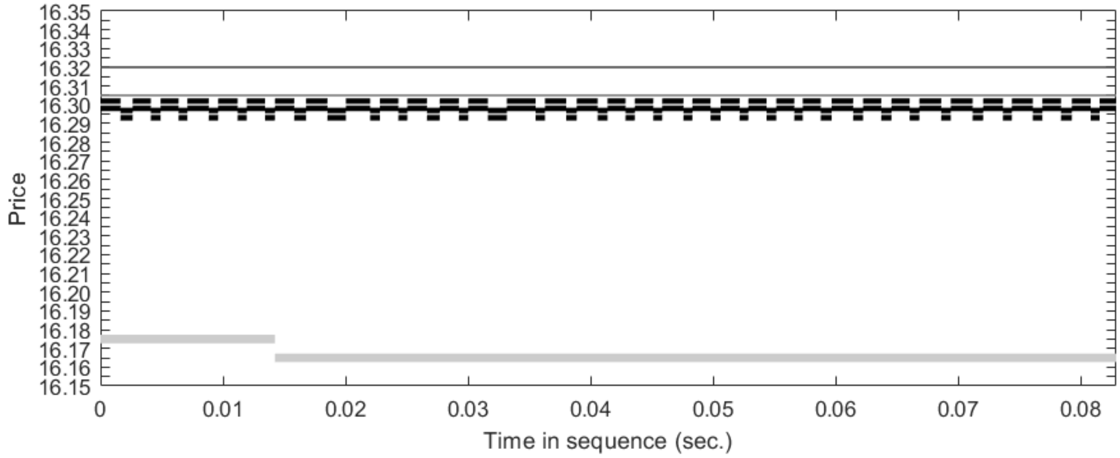


Panel C : 167 to 176 milliseconds in sequence

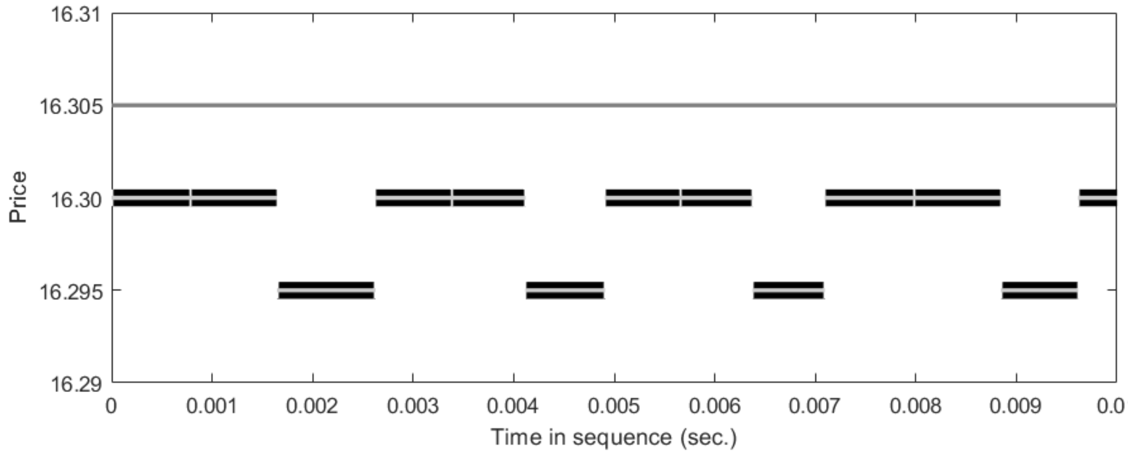


Example 2.1.1 – CEV stock on 2013-02-06 (12:32:30) – 101 ask orders

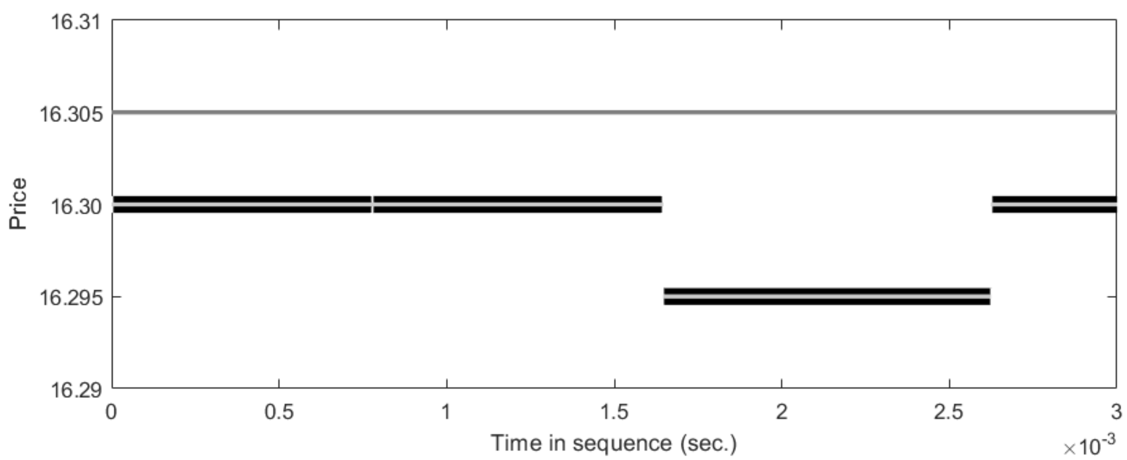
Panel A : Complete sequence – 82.9 milliseconds



Panel B : 0 to 10 milliseconds in sequence

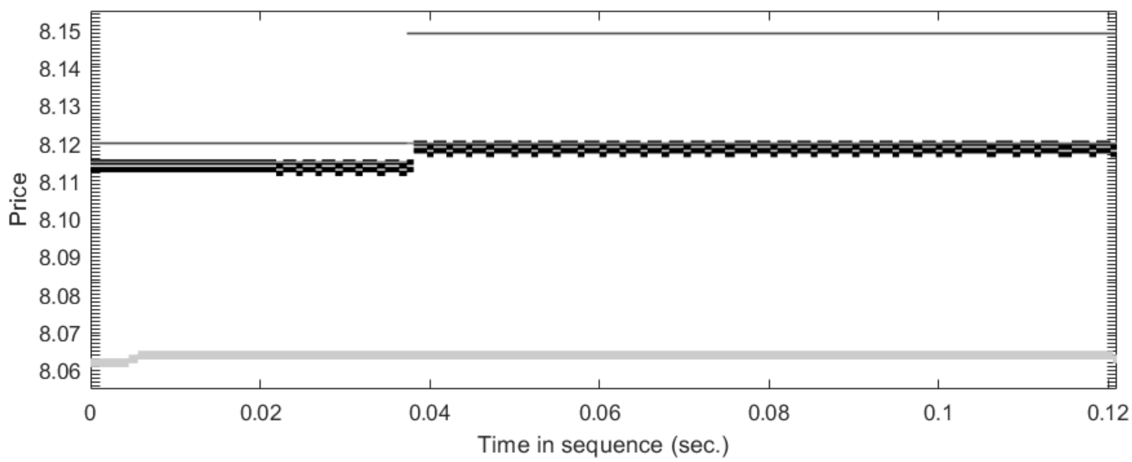


Panel C : 0 to 3 milliseconds in sequence

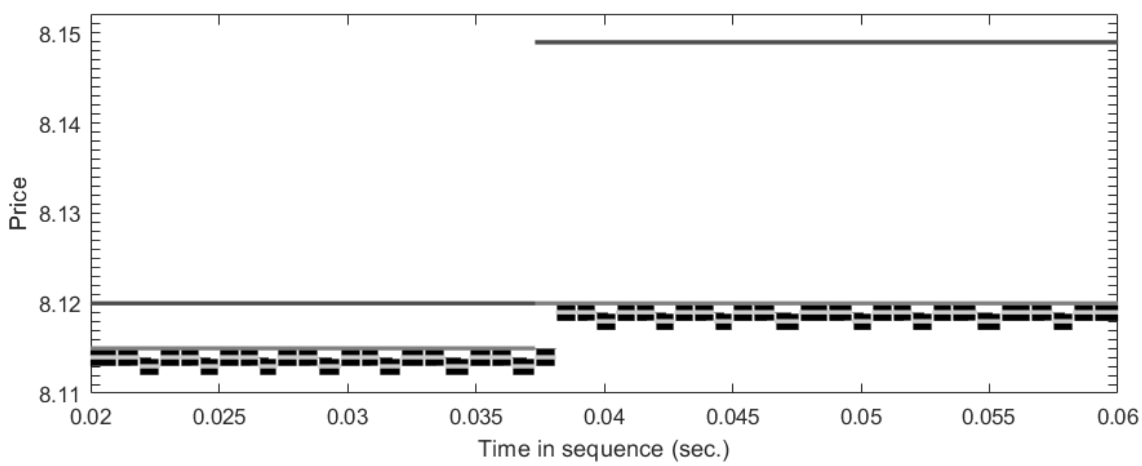


Example 2.1.13 – COM stock on 2013-04-11 (10:01:32) – 119 ask orders

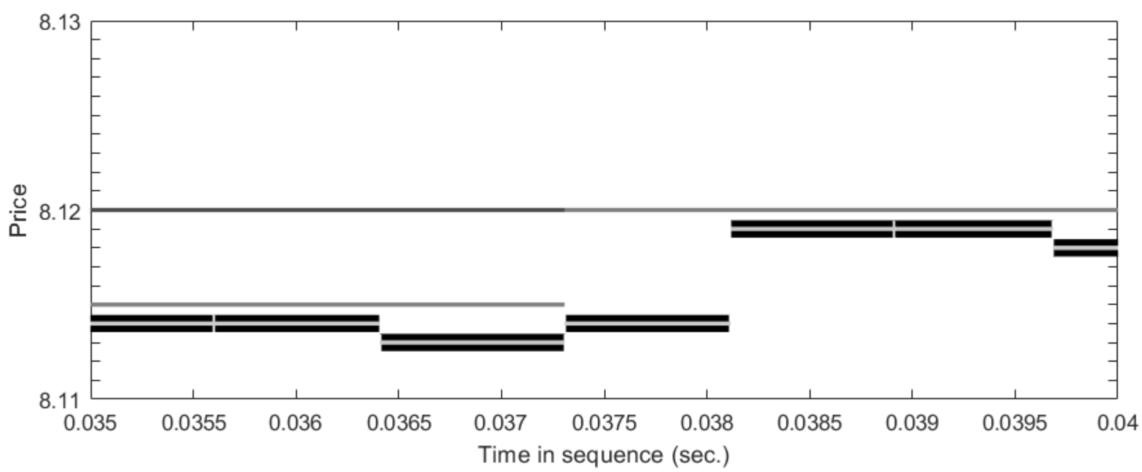
Panel A : Complete sequence – 121 milliseconds



Panel B : 20 to 60 milliseconds in sequence

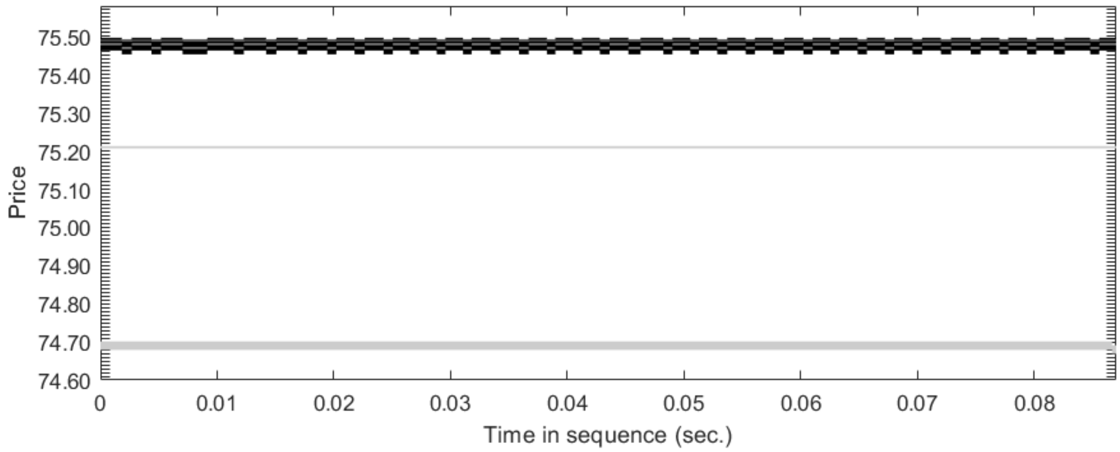


Panel C : 35 to 40 milliseconds in sequence

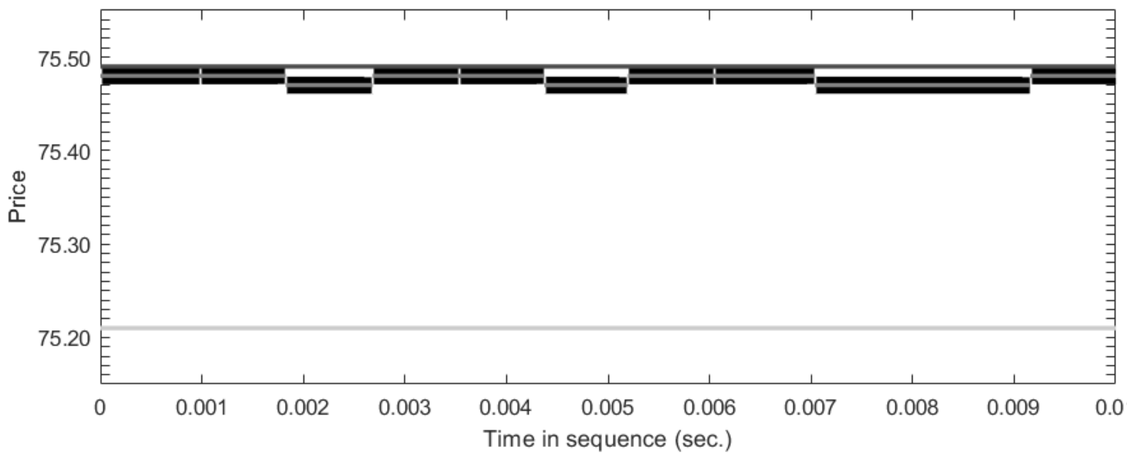


Example 2.2 – GSC1 stock on 2013-04-16 (9:31:23) – 104 ask orders

Panel A : Complete sequence – 87 milliseconds

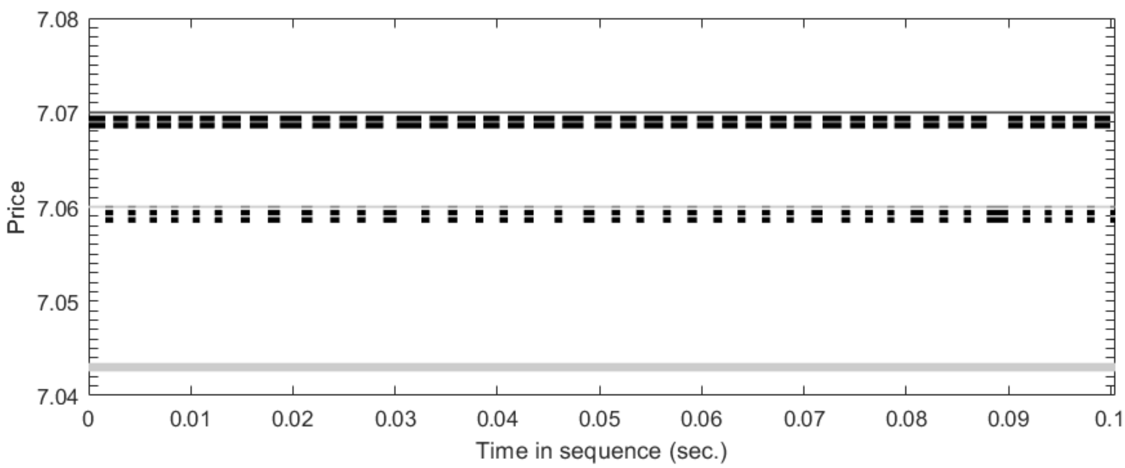


Panel B : 0 to 10 milliseconds in sequence

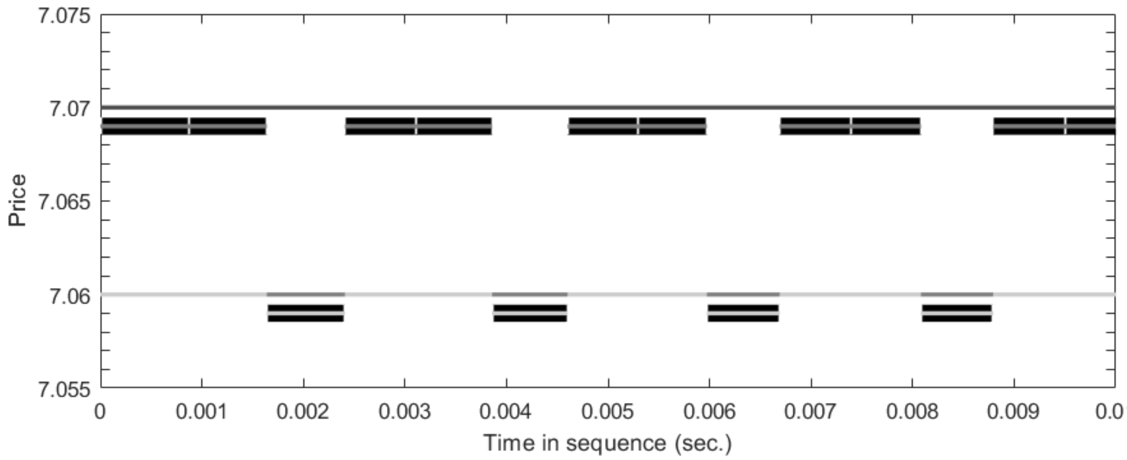


Example 2.4 – HAB stock on 2013-03-21 (17:15:34) – 120 ask orders

Panel A : Complete sequence – 100 milliseconds

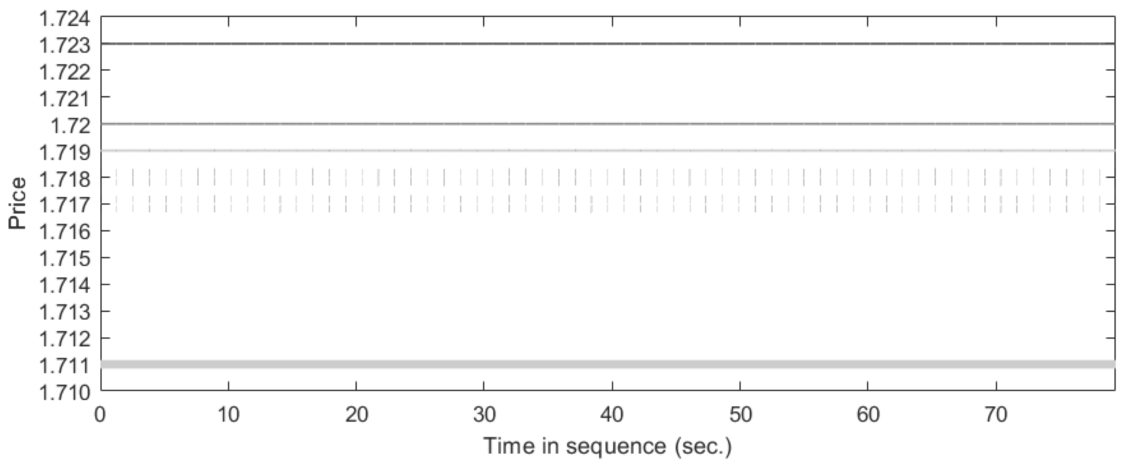


Panel B : 0 to 10 milliseconds in sequence

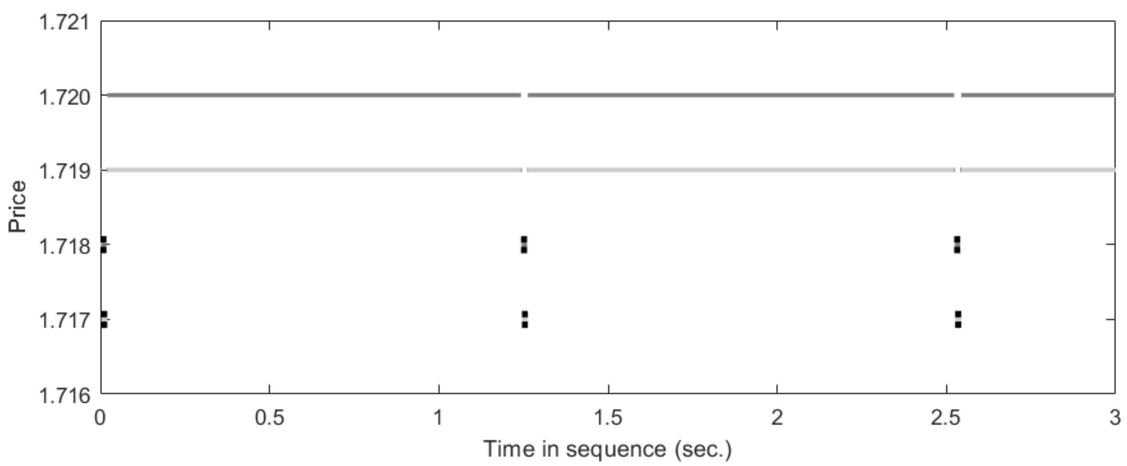


Example 2.12.1 – HDD stock on 2013-04-12 (11:03:37) – 126 ask orders

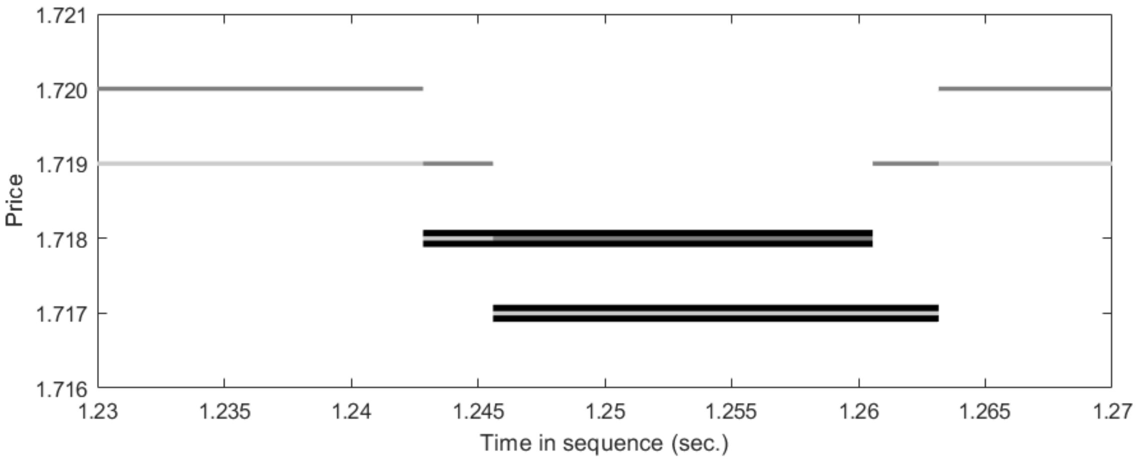
Panel A : Complete sequence – 1 minute 19 seconds



Panel B : 0 to 3 seconds in sequence

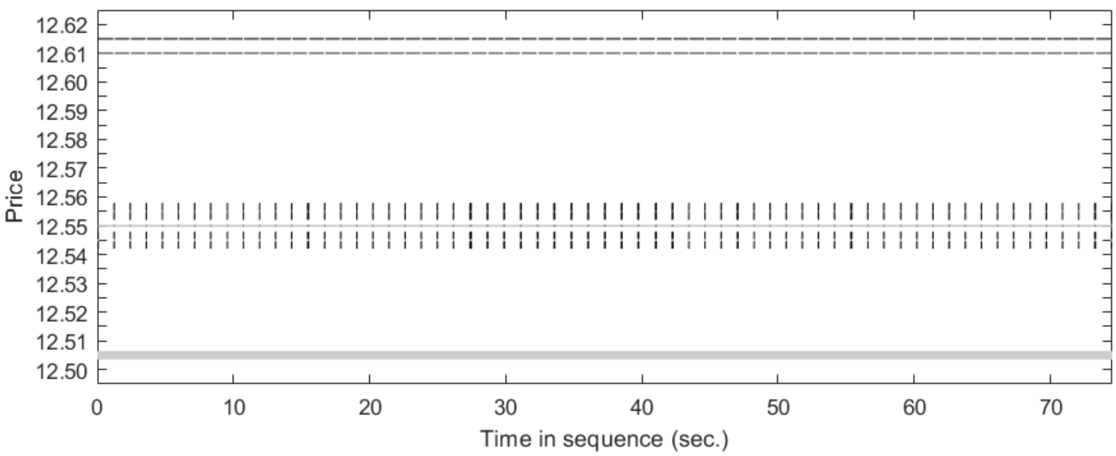


Panel C : 1.23 to 1.27 seconds in sequence

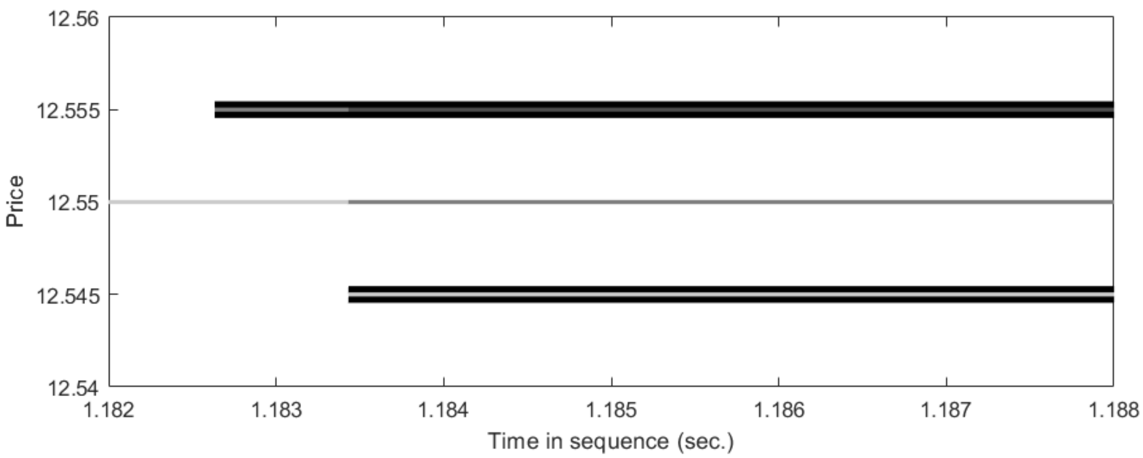


Example 2.13 – TTK stock on 2013-03-27 (15:19:20) – 125 ask orders

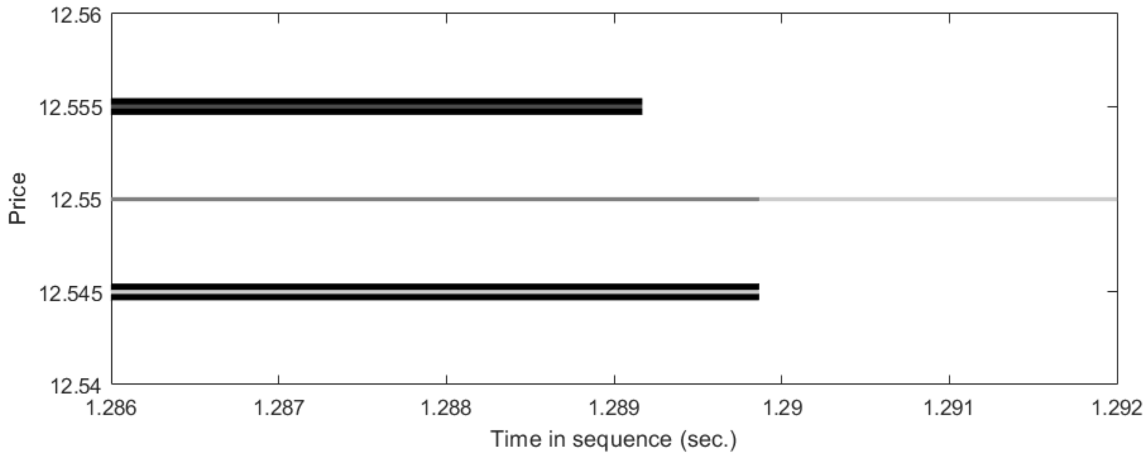
Panel A : Complete sequence – 1 minute 14 seconds



Panel B : 1182 milliseconds to 1188 milliseconds in sequence

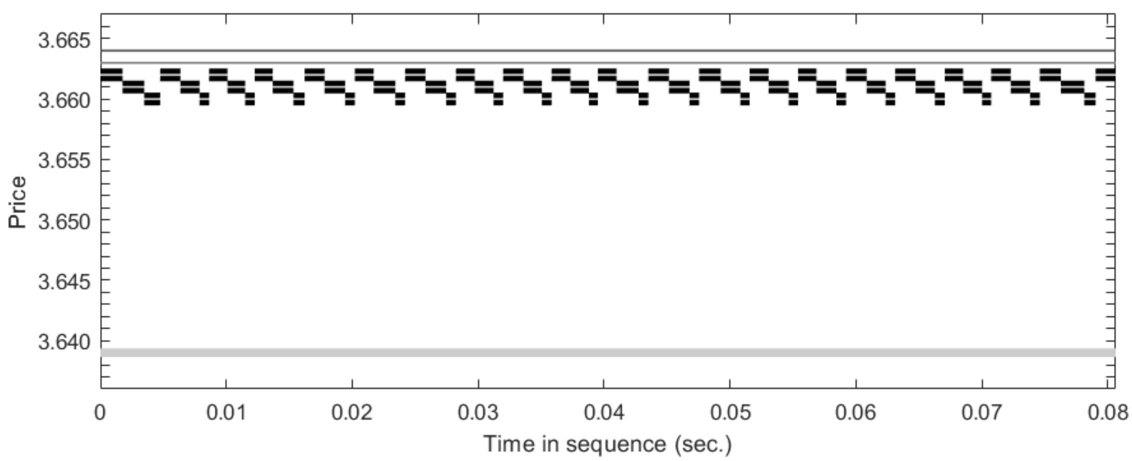


Panel C : 1286 milliseconds to 1292 milliseconds in sequence

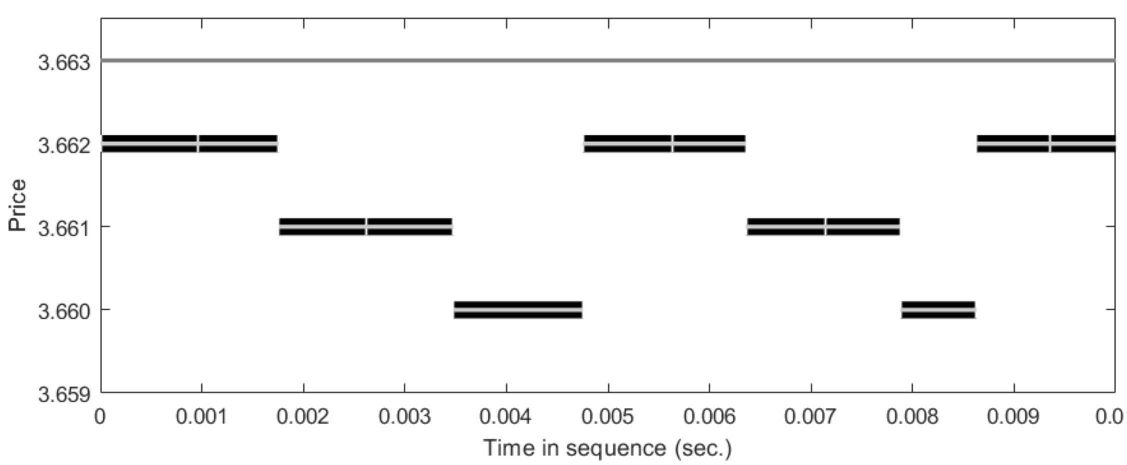


Example 3.1.1 – PMO stock on 2013-03-18 (12:25:48) – 102 ask orders

Panel A : Complete sequence – 81 milliseconds

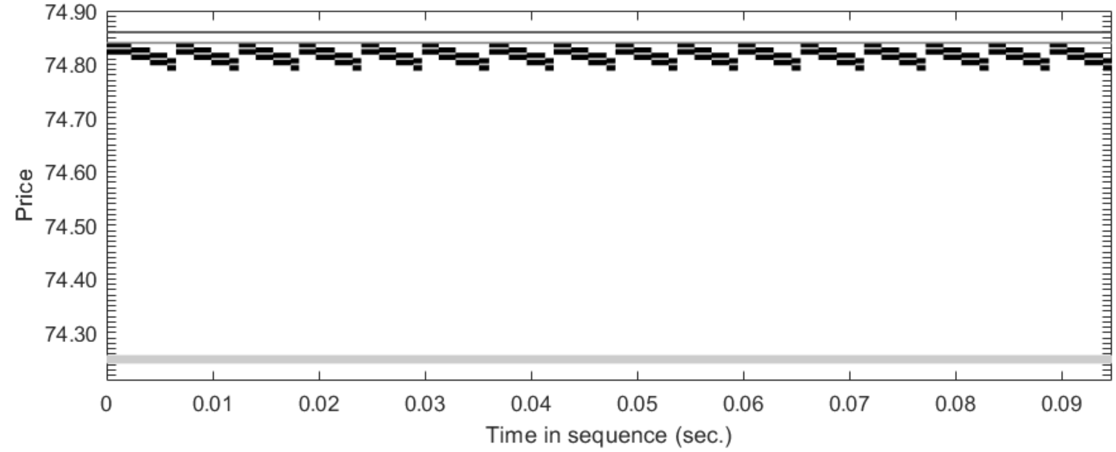


Panel B : 0 to 10 milliseconds in sequence

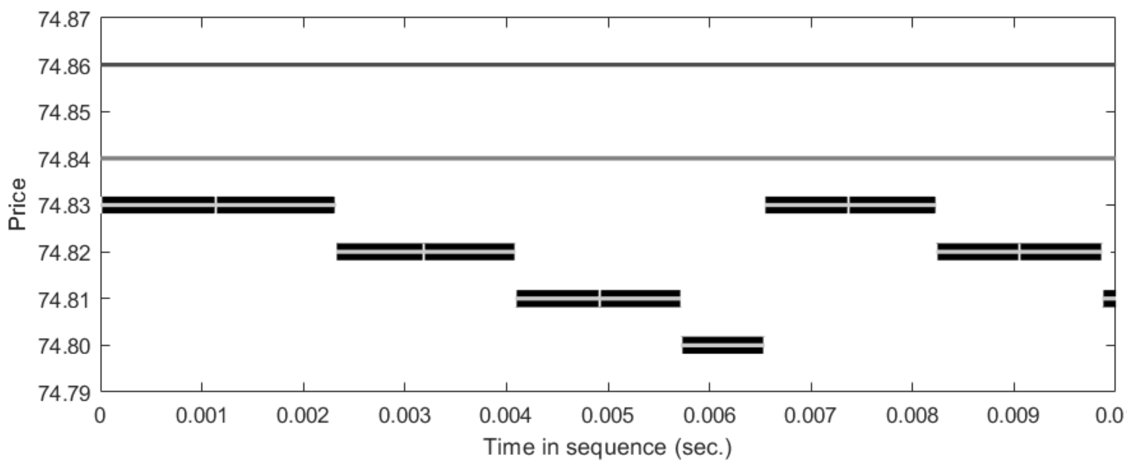


Example 4.1.1 – GSC1 stock on 2013-04-05 (13:51:47) – 112 ask orders

Panel A : Complete sequence – 95 milliseconds

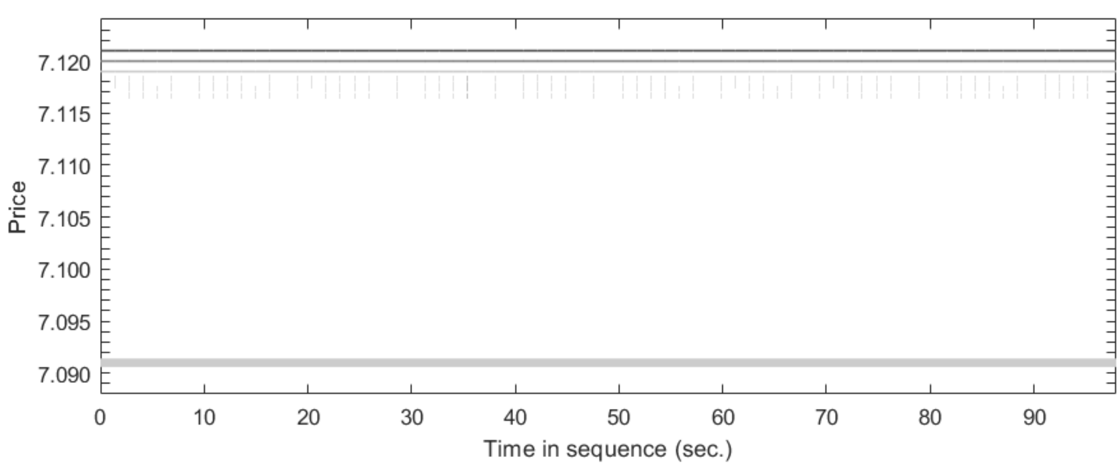


Panel B : 0 to 10 milliseconds in sequence

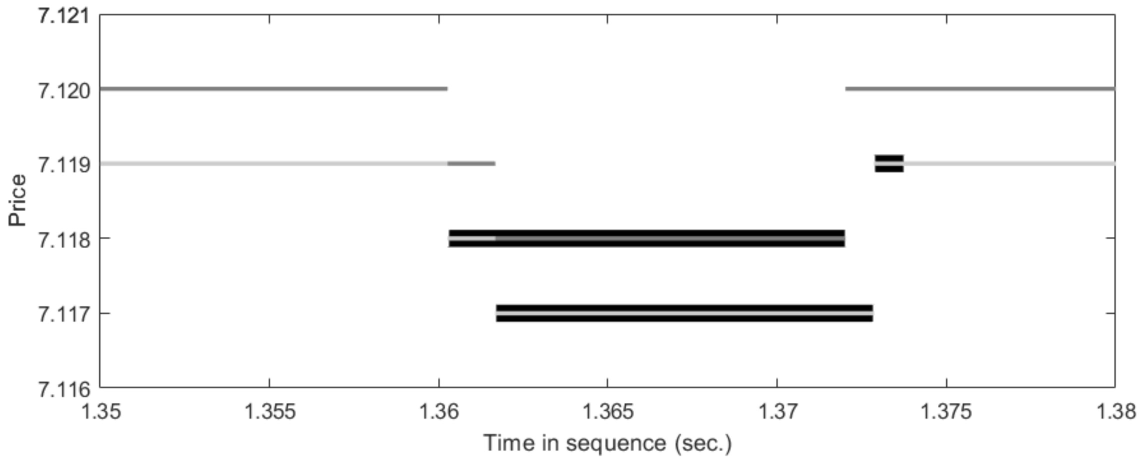


Example 3.4 – HAB stock on 2013-03-28 (14:02:33) – 219 ask orders

Panel A : Complete sequence – 98 seconds

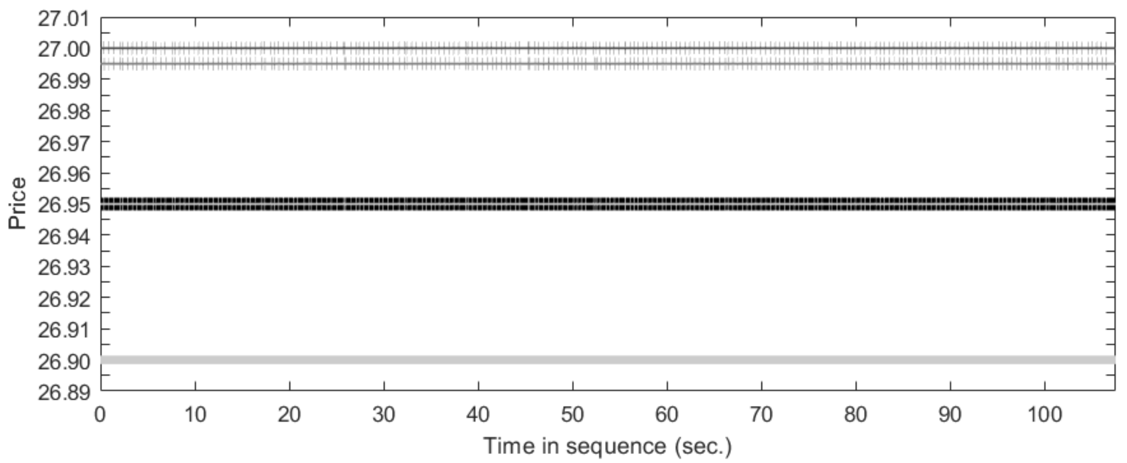


Panel B : 1.35 to 1.38 seconds in sequence

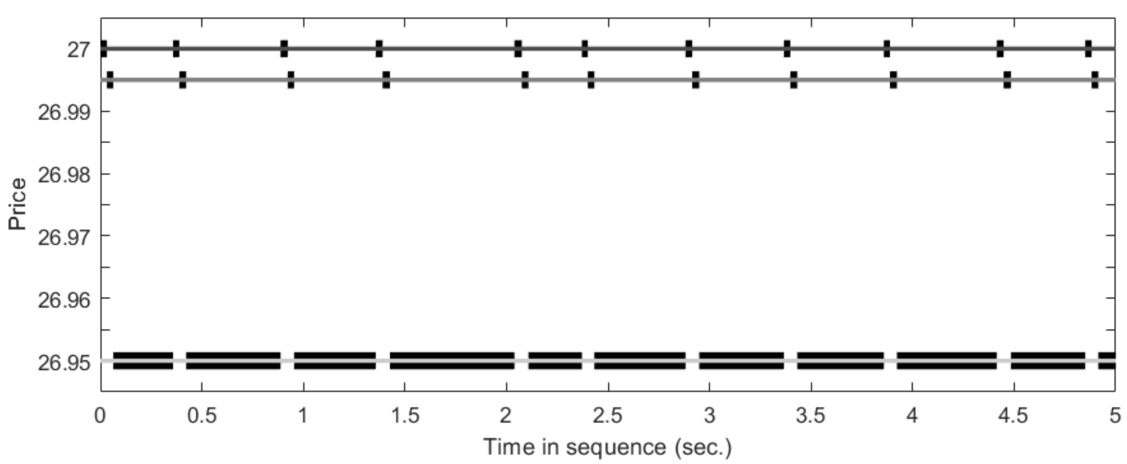


Example 3.5 – EVD stock on 2013-03-13 (15:10:04) – 643 ask orders

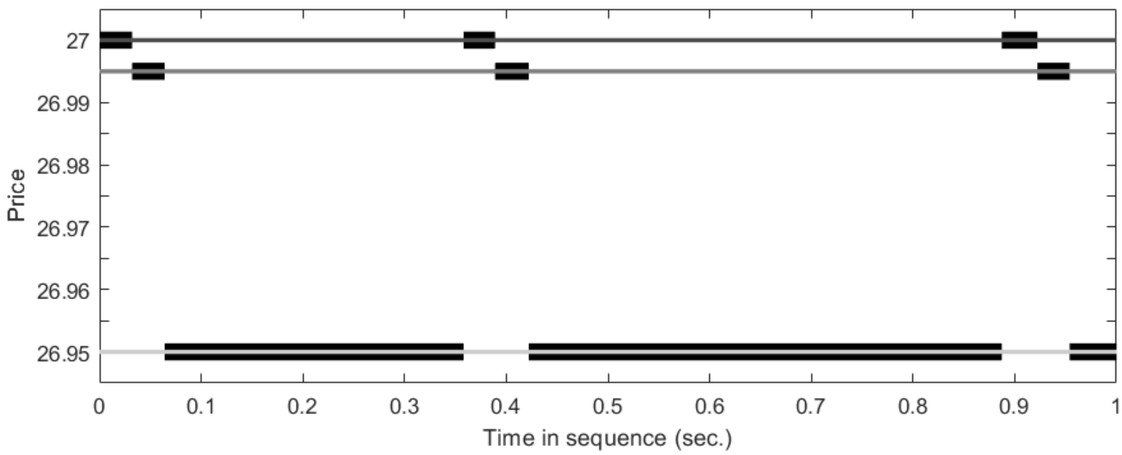
Panel A : Complete sequence – 108 seconds



Panel B : 0 to 5 seconds in sequence

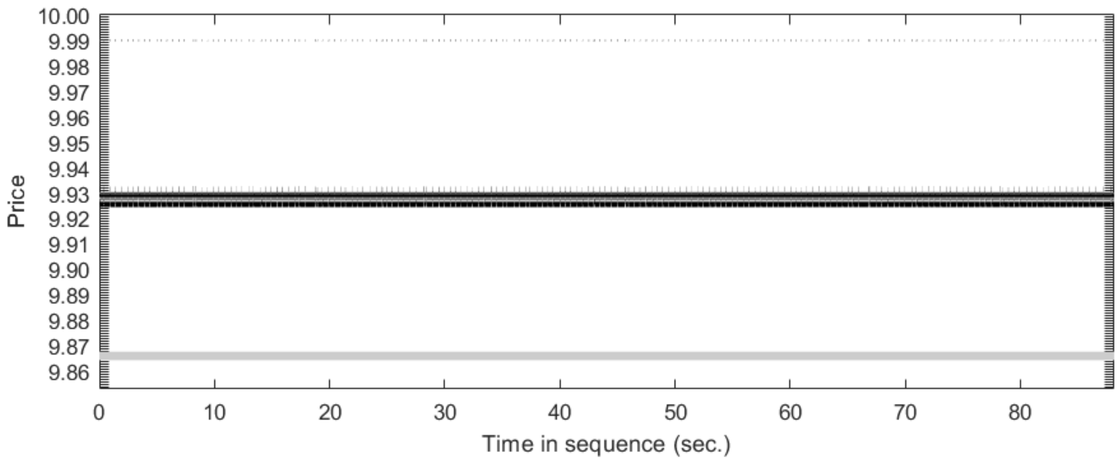


Panel C : 0 to 1 second in sequence

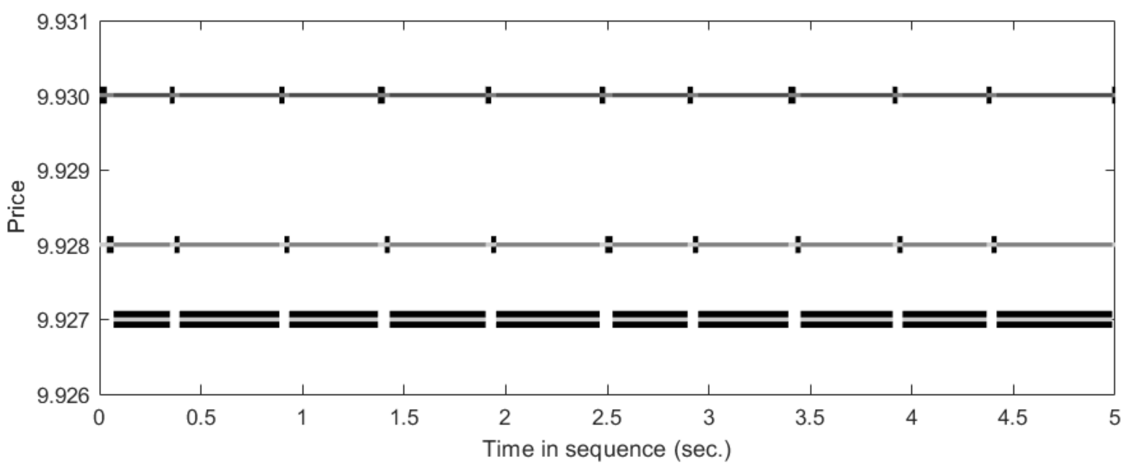


Example 3.6 – 2HR stock on 2013-04-15 (14:05:00) – 529 ask orders

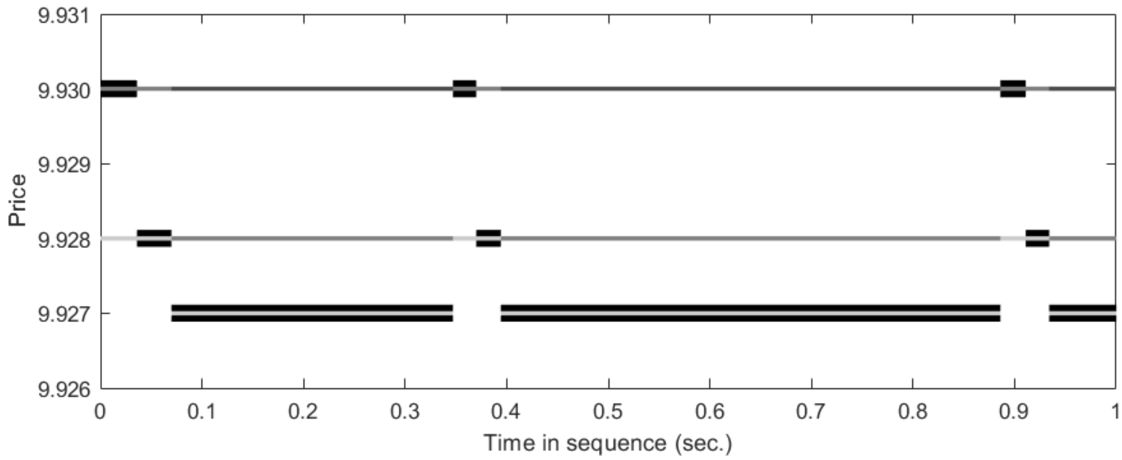
Panel A : Complete sequence – 88 seconds



Panel B : 0 to 5 seconds in sequence

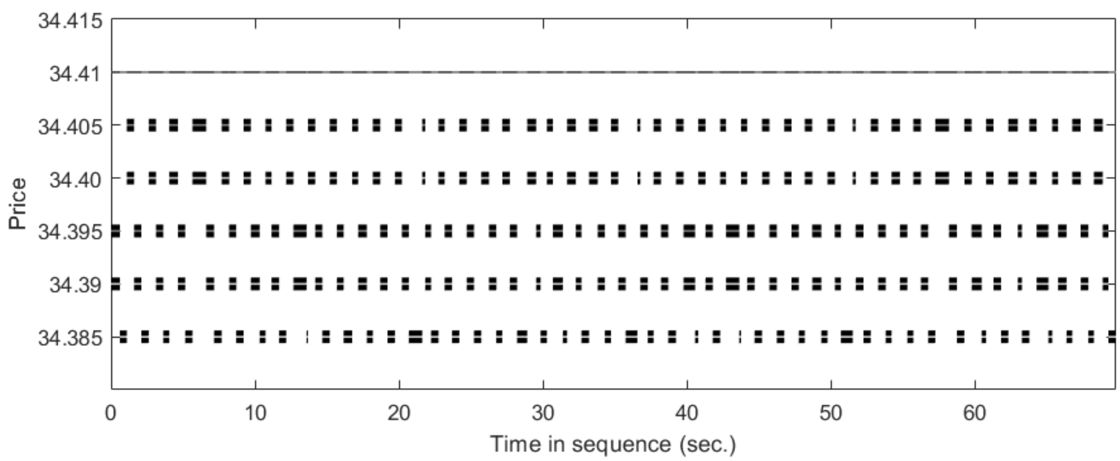


Panel C : 0 to 1 second in sequence

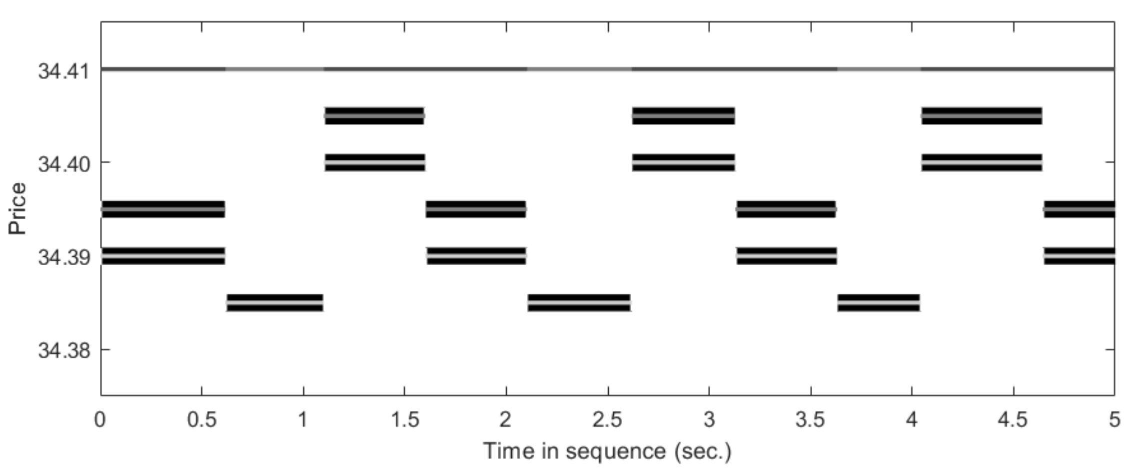


Example 5.1 – SLT stock on 2013-02-05 (17:10:38) – 228 ask orders

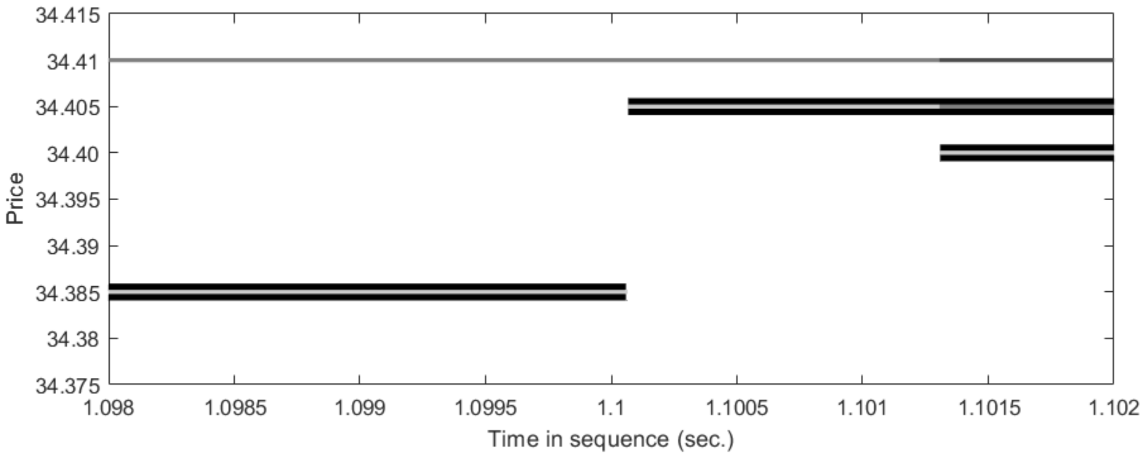
Panel A : Complete sequence – 69 seconds



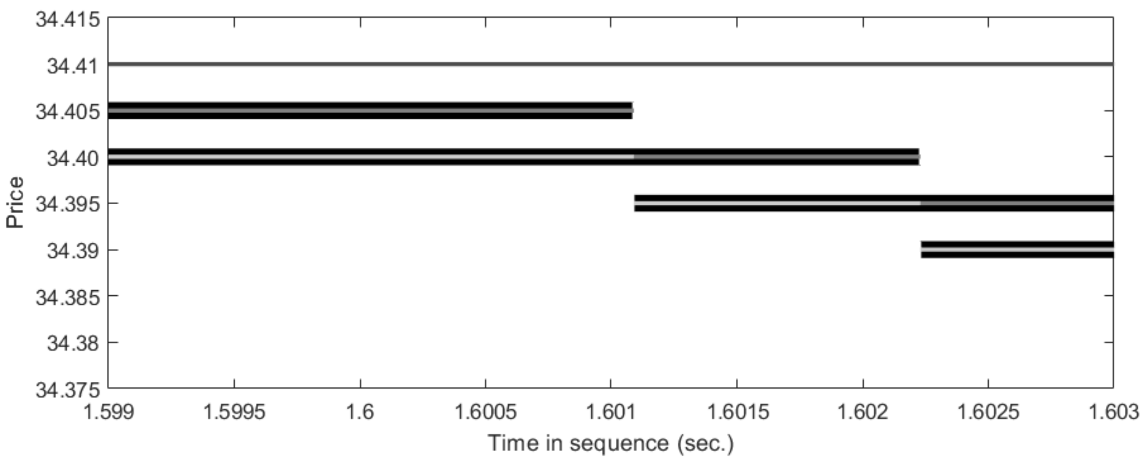
Panel B : 0 to 5 seconds in sequence



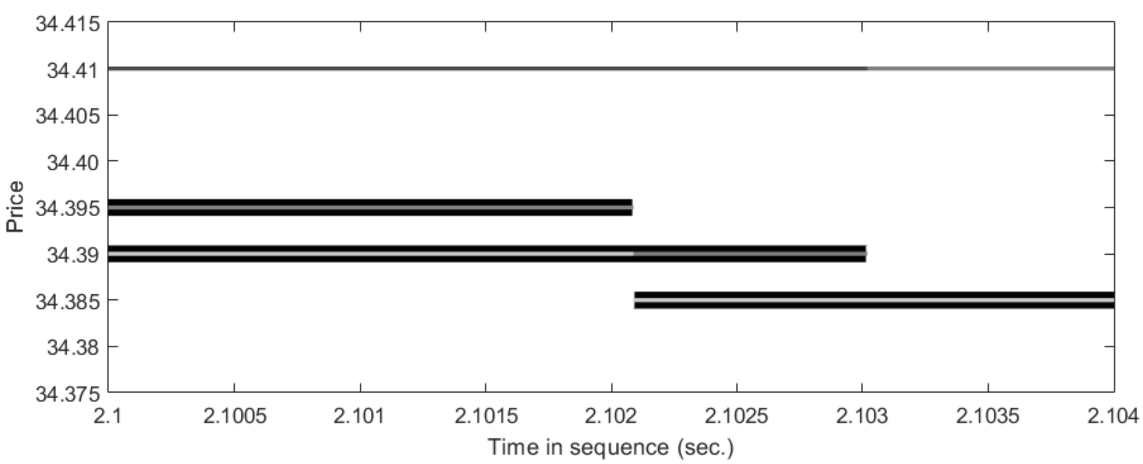
Panel C : 1.098 to 1.102 second in sequence



Panel D : 1.599 to 1.603 second in sequence

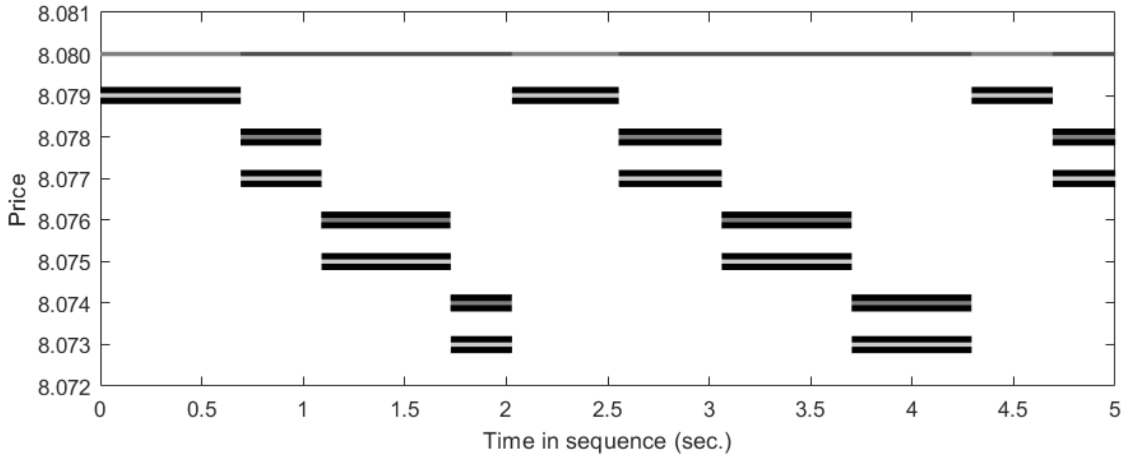


Panel E : 2.0 to 1.11 second in sequence



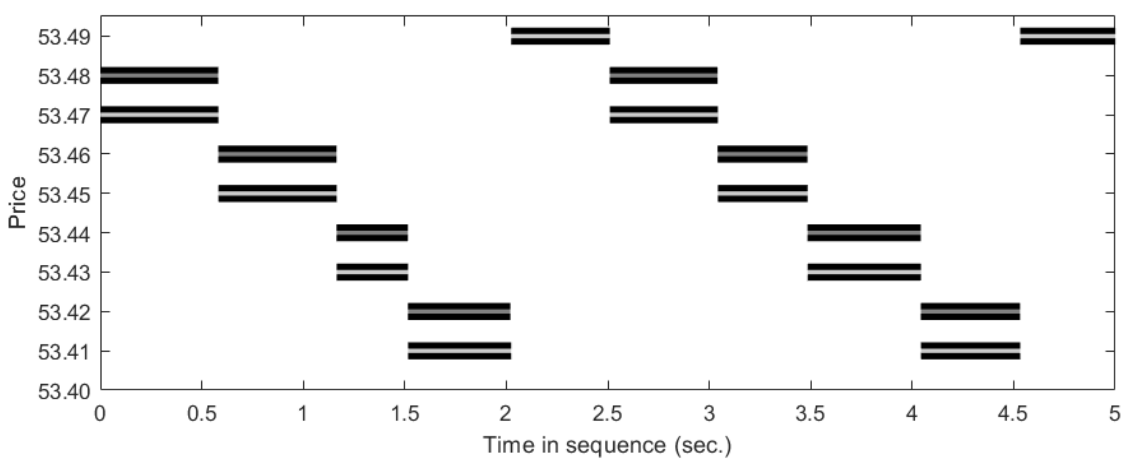
Example 7.1 – COM stock on 2013-04-09 (12:59:52) – 145 ask orders

Panel A : 0 to 5 seconds in sequence



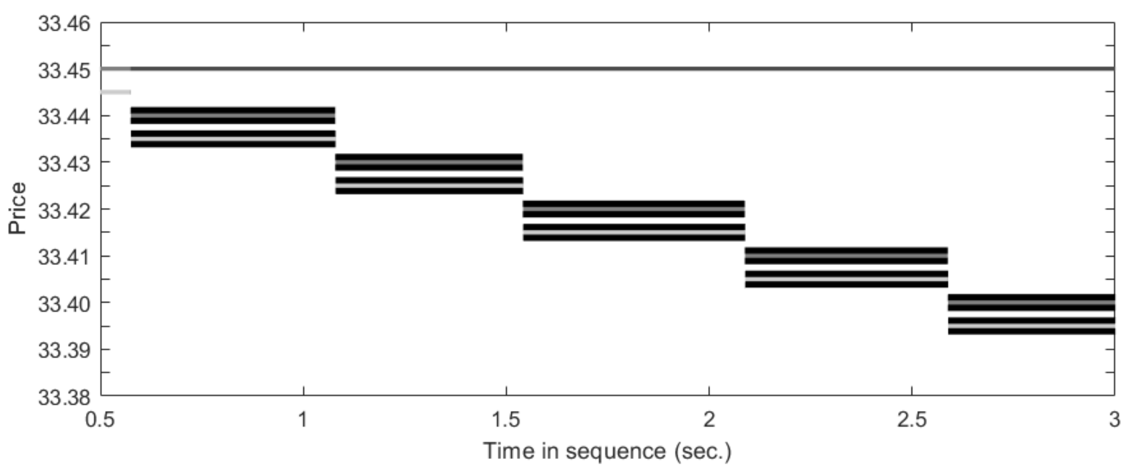
Example 9.1 – BIO3 stock on 2013-03-14 (13:04:13) – 170 ask orders

Panel A : 0 to 5 seconds in sequence



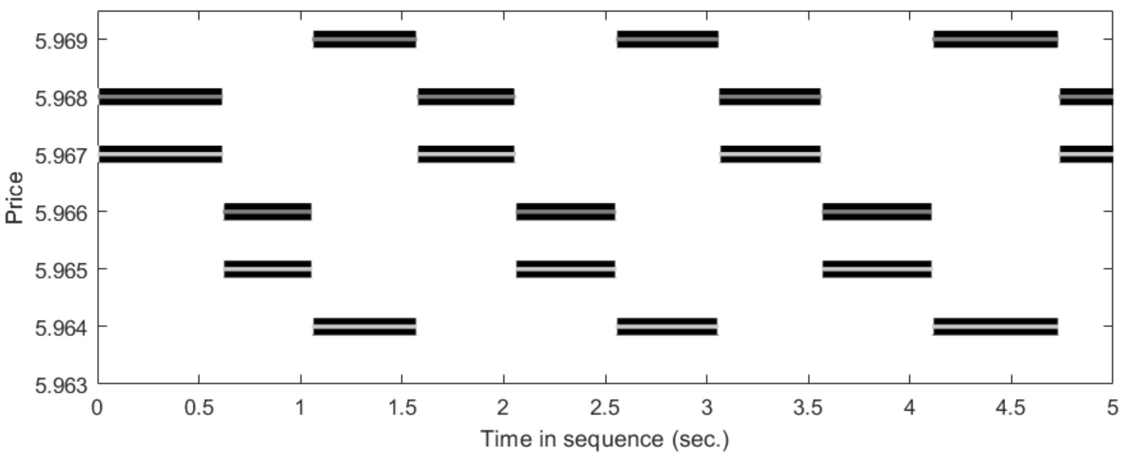
Example 10.1 – SLT stock on 2013-03-25 (15:34:43) – 114 ask orders

Panel A : 0.5 to 3 seconds in sequence



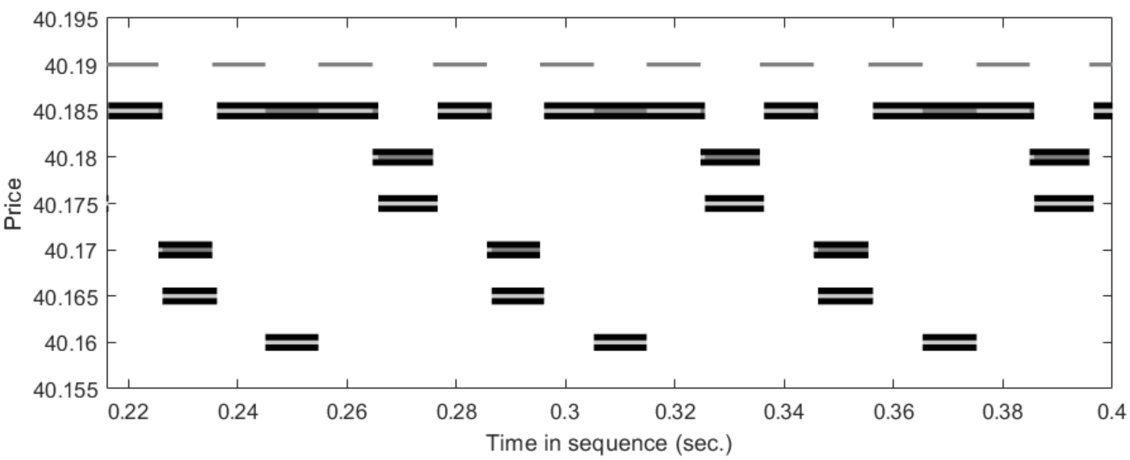
Example 6.1.1 – MLP stock on 2013-03-08 (10:53:19) – 488 ask orders

Panel A : 0 to 5 seconds in sequence

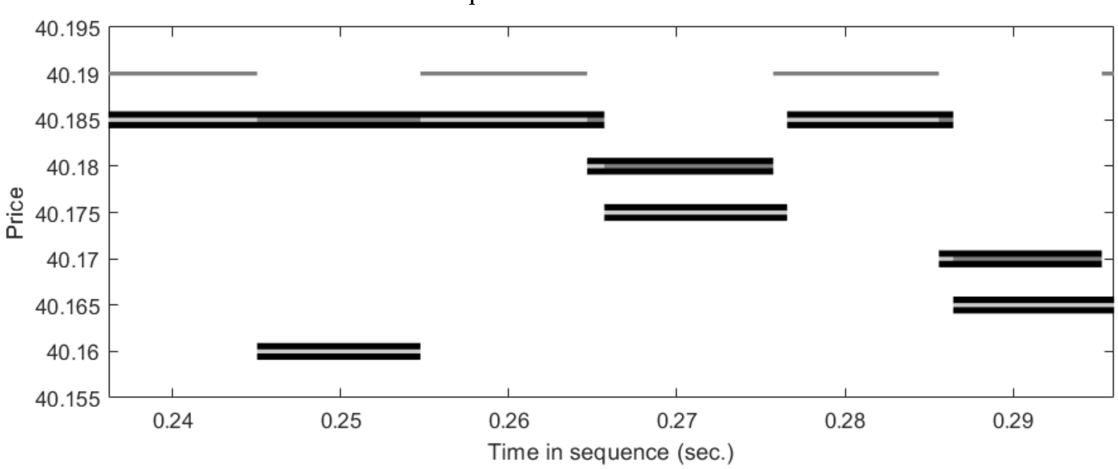


Example 6.3 – DEX stock on 2013-03-21 (13:03:48) – 207 ask orders

Panel A : 220 to 400 milliseconds in sequence



Panel B : 236 to 296 milliseconds in sequence



Example 8.1 – KWS stock on 2013-04-24 (9:42:38) – 170 bid orders

Panel A : 407 to 483 milliseconds in sequence

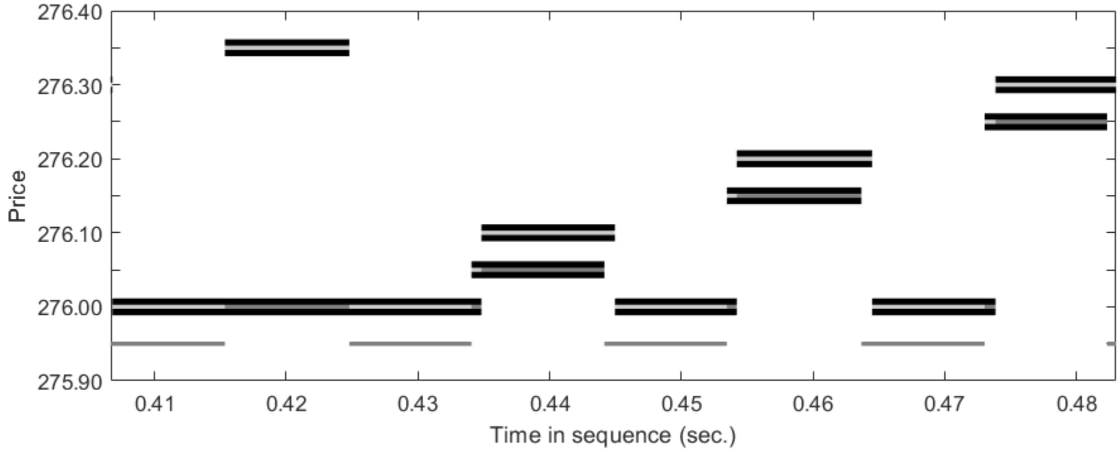
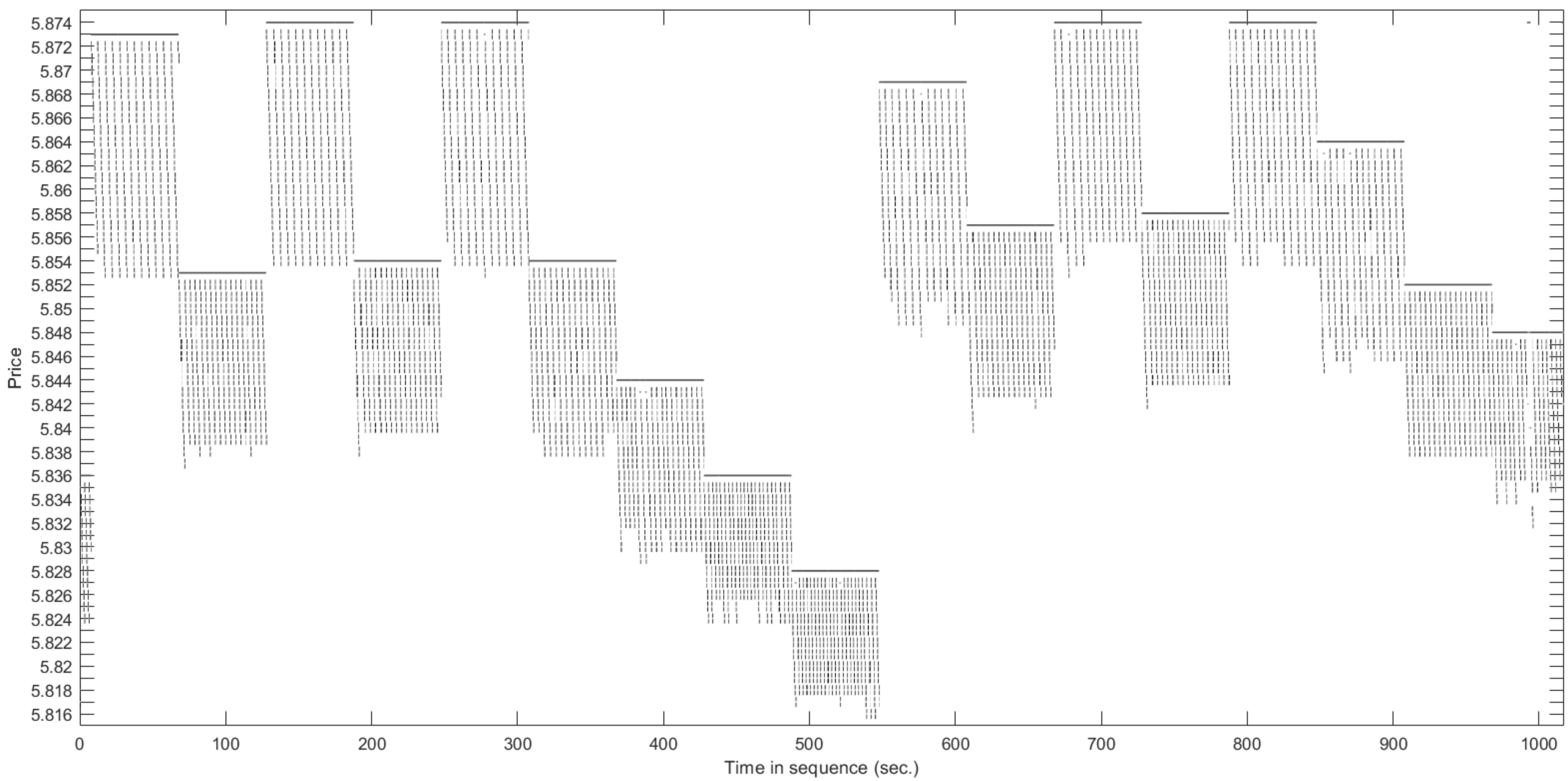


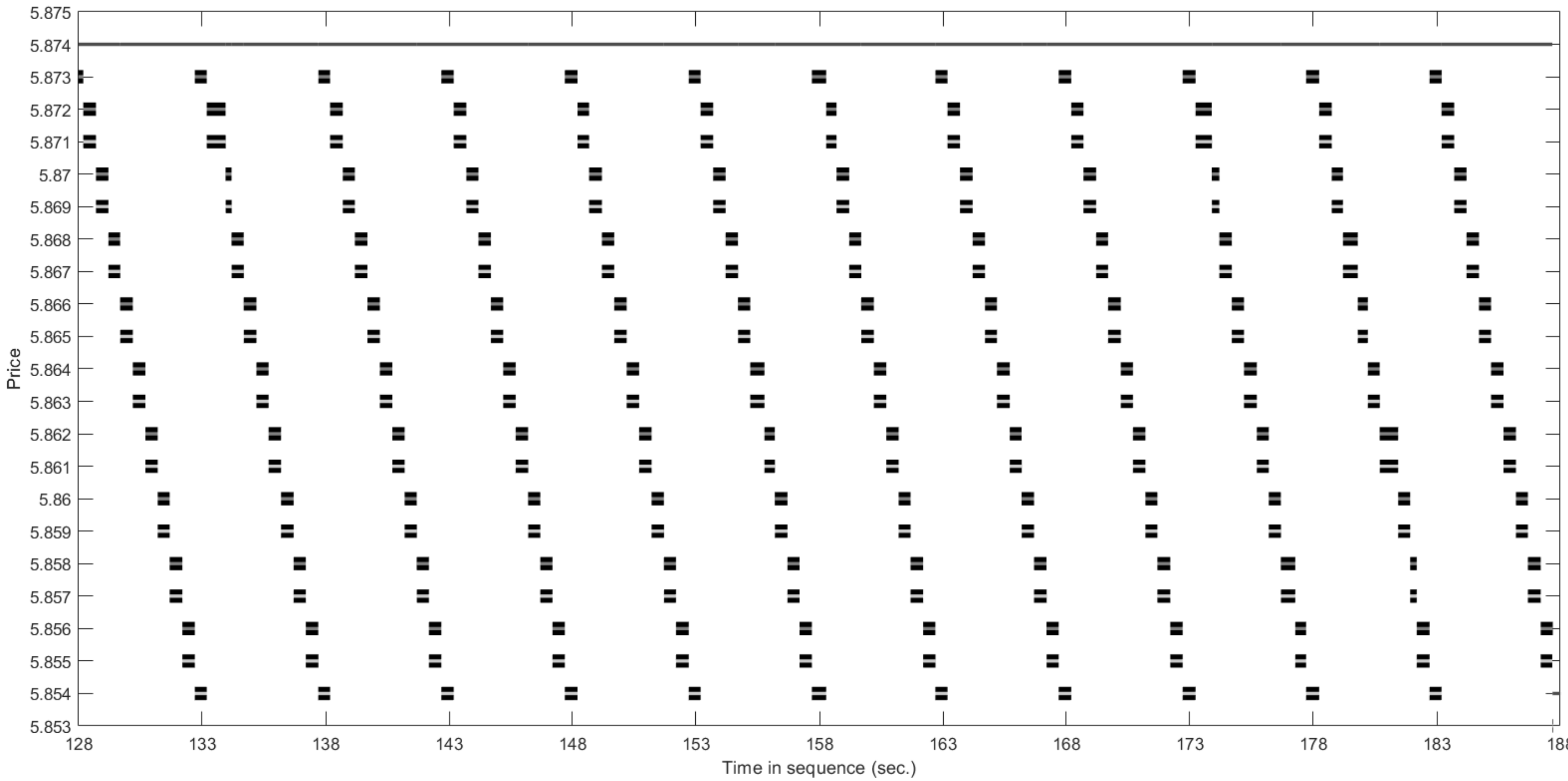
Figure 5.13 Long algorithmic orders sequence examples

Example 1 : MLP stock on 2013-03-13 (12:40:46) – 4093 ask orders

Panel A: Complete sequence – 16 minutes 57 seconds

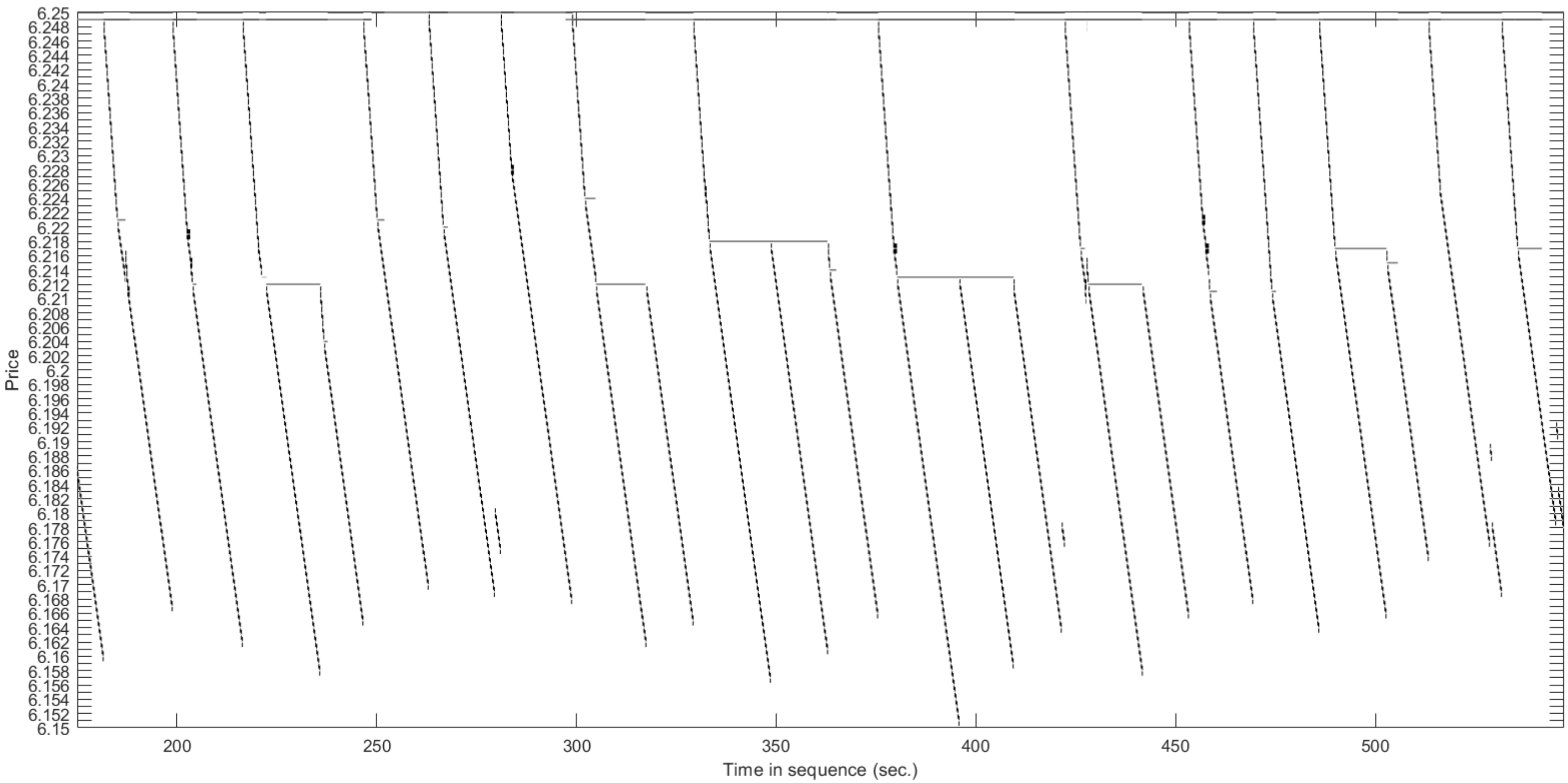


Panel B: 128 to 188 seconds in sequence



Example 2 : MLP stock on 2013-02-18 (9:35:34) – 2507 ask orders

Panel A: Complete sequence – 9 minutes 7 seconds



Panel B: 283 to 299 seconds in sequence

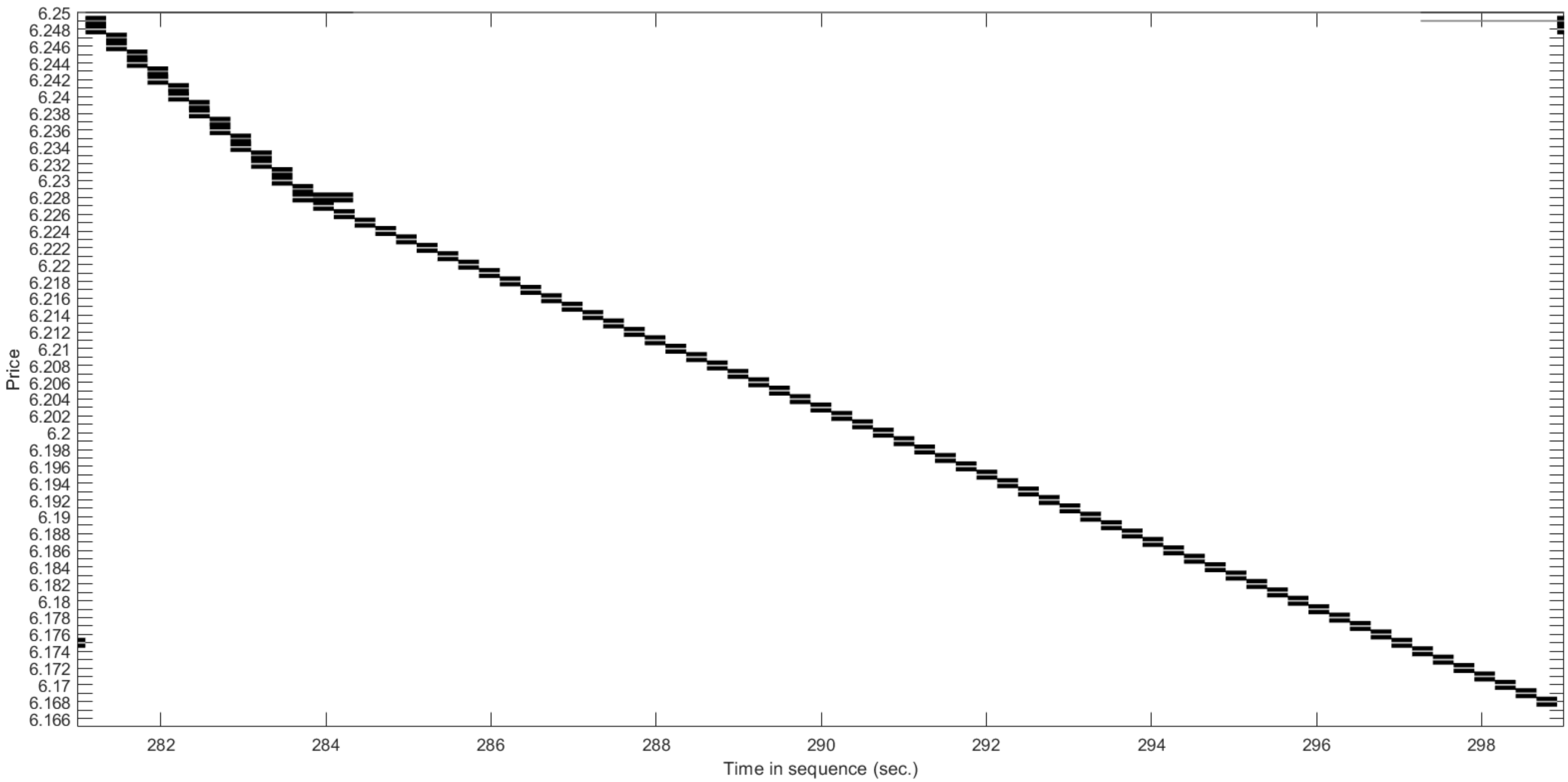
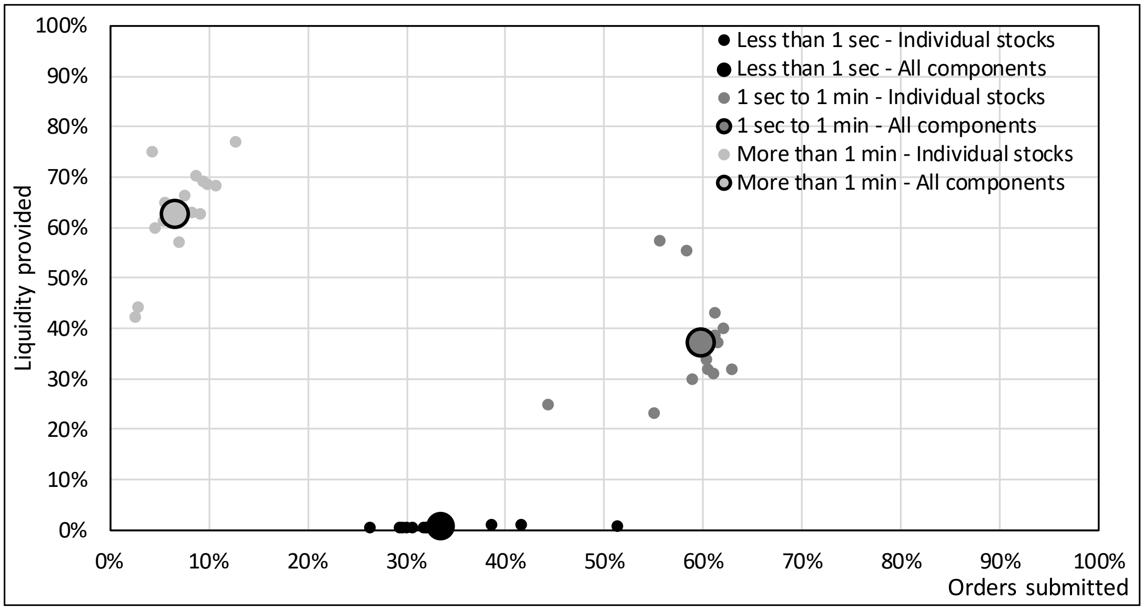
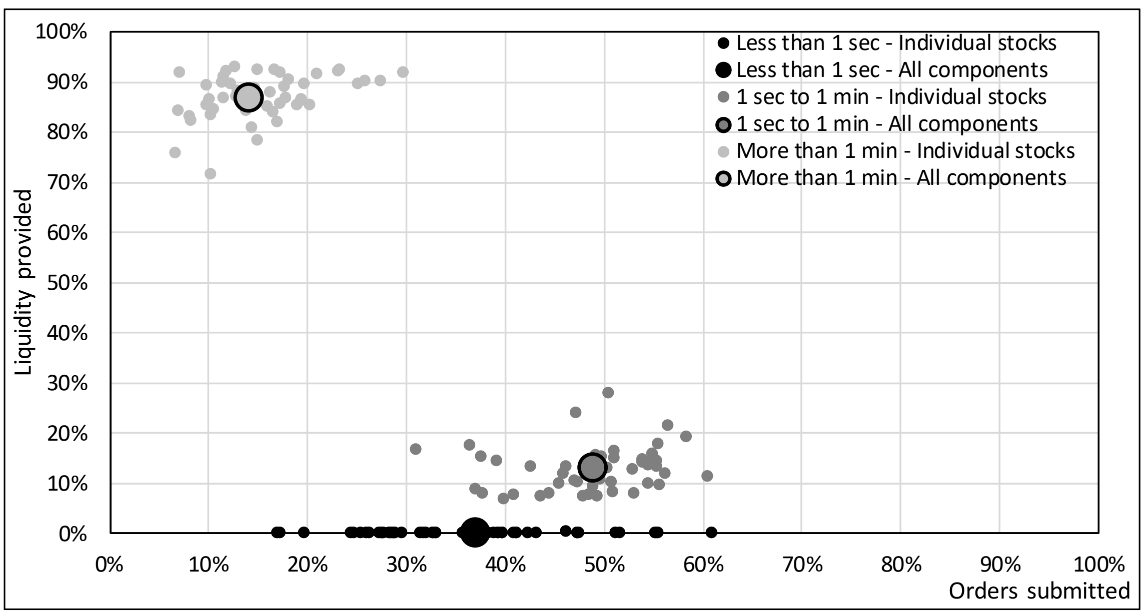


Figure 5.14 Liquidity provided vs number of orders proportions by orders category

Panel I: *DAX15* components



Panel II: MDAX components



Panel III: SDAX components

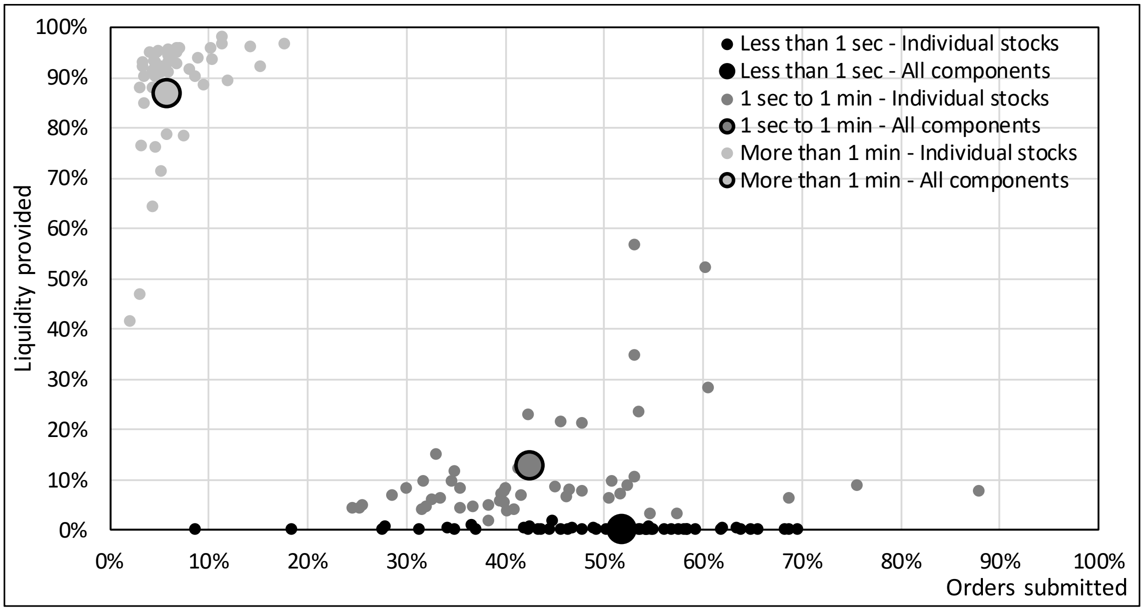


Table 5.1 Passive orders termination context

Panel I: *DAX15* components

		Price level of submission																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Price level of cancellation	1	57.7%	19.1%	6.7%	2.7%	1.8%	1.0%	0.9%	0.7%	0.5%	0.3%	0.2%	0.2%	0.1%	0.2%	0.2%	0.1%	0.1%	0.0%	0.0%	0.0%
	2	14.9%	52.1%	28.4%	10.0%	4.5%	2.7%	2.1%	1.7%	1.4%	0.9%	0.9%	0.7%	0.5%	0.6%	0.6%	0.5%	0.3%	0.1%	0.0%	0.0%
	3	4.0%	11.9%	31.2%	22.8%	8.3%	4.6%	3.2%	2.3%	1.9%	1.5%	1.7%	1.1%	0.8%	0.7%	0.7%	0.6%	0.5%	0.3%	0.1%	0.0%
	4	1.6%	4.8%	16.5%	29.7%	14.3%	8.3%	5.7%	3.6%	2.8%	2.1%	2.4%	1.8%	1.2%	0.9%	0.7%	0.7%	0.6%	0.4%	0.2%	0.1%
	5	0.8%	2.1%	5.2%	10.0%	35.2%	13.4%	7.7%	5.4%	4.2%	2.5%	2.9%	2.2%	1.5%	0.8%	0.6%	0.5%	0.5%	0.4%	0.4%	0.2%
	6	0.5%	1.7%	3.7%	12.0%	13.1%	35.1%	12.7%	7.3%	5.7%	3.3%	3.5%	3.0%	2.4%	1.4%	0.8%	0.6%	0.5%	0.3%	0.2%	0.1%
	7	0.6%	1.3%	2.5%	4.3%	8.1%	11.5%	29.3%	11.4%	6.9%	3.6%	4.3%	3.3%	2.3%	1.9%	1.2%	0.8%	0.6%	0.4%	0.2%	0.1%
	8	0.2%	0.6%	1.0%	1.9%	3.5%	6.6%	12.3%	25.2%	9.1%	4.2%	4.9%	3.9%	2.7%	1.8%	1.3%	1.2%	1.0%	0.6%	0.3%	0.2%
	9	0.1%	0.4%	0.6%	1.2%	2.4%	4.5%	7.9%	14.7%	17.5%	5.2%	6.0%	5.1%	3.5%	2.2%	1.4%	1.1%	1.0%	0.8%	0.4%	0.2%
	10	0.1%	0.3%	0.5%	1.0%	1.9%	3.3%	5.3%	8.5%	13.2%	12.4%	7.6%	6.4%	5.4%	3.2%	1.9%	1.3%	0.9%	0.6%	0.4%	0.2%
	11	0.1%	0.4%	0.7%	1.2%	2.1%	3.3%	5.3%	9.0%	21.1%	48.6%	20.0%	7.9%	5.9%	5.4%	3.0%	1.8%	1.4%	1.0%	0.5%	0.2%
	12	0.0%	0.1%	0.3%	0.6%	0.9%	1.2%	1.9%	3.1%	5.3%	6.4%	15.0%	18.0%	7.7%	5.0%	4.8%	2.8%	2.0%	1.4%	0.8%	0.3%
	13	0.0%	0.1%	0.2%	0.4%	0.8%	1.0%	1.4%	1.9%	3.0%	3.2%	10.8%	13.9%	17.8%	6.3%	4.0%	3.7%	2.9%	2.0%	1.1%	0.4%
	14	0.0%	0.1%	0.2%	0.4%	0.7%	1.0%	1.1%	1.3%	2.0%	1.9%	6.8%	10.2%	14.5%	16.4%	6.0%	4.0%	3.5%	3.1%	1.7%	0.5%
	15	0.0%	0.1%	0.1%	0.3%	0.5%	0.7%	0.8%	0.9%	1.3%	1.2%	4.6%	7.1%	10.1%	15.2%	18.1%	7.1%	4.3%	3.5%	2.4%	0.7%
	16	0.0%	0.0%	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%	0.8%	0.7%	2.6%	4.8%	7.7%	12.2%	17.4%	22.2%	8.3%	4.6%	2.8%	1.0%
	17	0.0%	0.0%	0.1%	0.1%	0.2%	0.2%	0.3%	0.4%	0.6%	0.5%	1.6%	3.1%	5.0%	9.8%	13.9%	18.5%	23.9%	9.3%	4.5%	1.3%
	18	0.0%	0.0%	0.1%	0.1%	0.1%	0.2%	0.2%	0.2%	0.4%	0.3%	1.0%	1.8%	3.3%	5.9%	9.9%	12.5%	16.7%	22.6%	9.6%	2.5%
	19	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.1%	0.2%	0.2%	0.2%	0.5%	1.0%	1.7%	3.1%	4.7%	7.1%	9.9%	15.1%	19.4%	5.1%
	20	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%	0.3%	0.4%	0.6%	0.9%	1.4%	2.1%	3.2%	4.9%	3.8%
Above 20		0.1%	0.1%	0.1%	0.1%	0.3%	0.4%	0.8%	1.4%	1.7%	1.0%	2.6%	4.1%	5.2%	6.3%	7.7%	11.3%	18.8%	30.6%	49.8%	83.0%
Executed		19.3%	4.8%	1.9%	1.1%	0.7%	0.4%	0.3%	0.3%	0.2%	0.1%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%

Panel II: MDAX index components

	Price level of submission																			
Price level of cancellation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	48.7%	13.0%	3.7%	2.3%	1.9%	1.3%	0.7%	0.5%	0.3%	0.3%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%
2	29.9%	52.2%	14.7%	7.0%	5.0%	3.3%	1.9%	1.2%	0.8%	0.6%	0.5%	0.4%	0.3%	0.3%	0.2%	0.2%	0.1%	0.1%	0.0%	0.0%
3	5.6%	18.5%	46.0%	21.7%	8.9%	5.1%	3.2%	1.9%	1.3%	1.0%	0.8%	0.6%	0.4%	0.3%	0.3%	0.2%	0.1%	0.1%	0.0%	0.0%
4	2.1%	6.3%	18.5%	33.1%	14.6%	8.1%	4.9%	3.1%	2.0%	1.4%	1.2%	0.9%	0.6%	0.4%	0.3%	0.2%	0.2%	0.1%	0.1%	0.0%
5	0.8%	2.5%	7.1%	13.8%	33.4%	14.3%	7.7%	4.7%	2.9%	2.1%	1.7%	1.2%	0.8%	0.6%	0.4%	0.3%	0.2%	0.1%	0.1%	0.0%
6	0.5%	1.5%	3.2%	12.4%	18.0%	31.6%	14.1%	7.4%	4.3%	2.9%	2.3%	1.7%	1.2%	0.8%	0.6%	0.4%	0.2%	0.2%	0.1%	0.0%
7	0.7%	1.1%	1.9%	3.8%	7.5%	16.3%	30.0%	13.6%	6.8%	4.3%	3.1%	2.1%	1.5%	1.1%	0.8%	0.5%	0.3%	0.2%	0.1%	0.1%
8	0.2%	0.7%	1.2%	1.9%	3.7%	8.0%	16.8%	29.1%	12.8%	7.4%	4.7%	3.1%	2.1%	1.4%	1.0%	0.7%	0.5%	0.3%	0.2%	0.1%
9	0.1%	0.4%	0.7%	1.0%	2.1%	4.1%	8.2%	17.0%	27.9%	17.5%	8.3%	4.8%	3.1%	2.0%	1.3%	0.9%	0.6%	0.4%	0.2%	0.1%
10	0.1%	0.2%	0.4%	0.6%	1.2%	2.4%	4.5%	9.2%	20.5%	28.0%	12.0%	6.4%	3.9%	2.6%	1.7%	1.2%	0.8%	0.5%	0.3%	0.1%
11	0.0%	0.2%	0.3%	0.4%	0.7%	1.3%	2.4%	4.4%	8.4%	14.0%	28.6%	11.7%	5.9%	3.6%	2.3%	1.6%	1.1%	0.7%	0.4%	0.2%
12	0.0%	0.2%	0.3%	0.3%	0.5%	0.9%	1.4%	2.4%	4.2%	7.9%	14.6%	30.7%	11.5%	5.4%	3.3%	2.2%	1.5%	1.0%	0.6%	0.2%
13	0.0%	0.1%	0.3%	0.3%	0.4%	0.6%	0.9%	1.5%	2.4%	4.4%	8.3%	14.7%	33.4%	11.5%	5.2%	3.1%	2.0%	1.3%	0.8%	0.3%
14	0.0%	0.1%	0.2%	0.2%	0.3%	0.5%	0.7%	1.0%	1.5%	2.6%	4.8%	8.2%	14.3%	35.9%	11.2%	4.9%	2.8%	1.8%	1.0%	0.4%
15	0.0%	0.0%	0.1%	0.1%	0.2%	0.3%	0.5%	0.7%	0.9%	1.6%	2.9%	4.6%	7.7%	14.0%	38.1%	11.0%	4.6%	2.6%	1.4%	0.5%
16	0.0%	0.0%	0.1%	0.1%	0.2%	0.2%	0.3%	0.5%	0.6%	1.0%	1.7%	2.7%	4.4%	7.3%	13.6%	39.7%	10.8%	4.4%	2.2%	0.7%
17	0.0%	0.0%	0.0%	0.1%	0.1%	0.2%	0.2%	0.3%	0.4%	0.6%	1.0%	1.5%	2.4%	4.0%	6.8%	13.3%	40.6%	10.5%	3.9%	1.2%
18	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.2%	0.2%	0.4%	0.6%	0.9%	1.3%	2.1%	3.5%	6.1%	12.8%	39.4%	9.5%	2.2%
19	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.1%	0.1%	0.2%	0.3%	0.4%	0.7%	1.0%	1.7%	2.8%	5.0%	11.2%	35.7%	6.0%
20	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.1%	0.2%	0.3%	0.6%	1.4%	4.8%
Above 20	0.0%	0.1%	0.1%	0.2%	0.4%	0.7%	1.0%	1.2%	1.3%	1.7%	2.3%	3.1%	4.1%	5.5%	7.5%	10.6%	15.4%	24.6%	41.9%	83.0%
Executed	11.1%	2.8%	1.2%	0.7%	0.6%	0.5%	0.3%	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%

Panel III: SDAX index components

	Price level of submission																			
Price level of cancellation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	31.1%	5.0%	1.2%	0.4%	0.2%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
2	54.9%	50.2%	9.2%	2.1%	0.7%	0.3%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
3	9.0%	37.5%	50.4%	8.5%	2.3%	0.9%	0.4%	0.2%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
4	1.7%	4.5%	30.8%	64.3%	10.1%	2.5%	1.1%	0.5%	0.3%	0.2%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
5	0.8%	1.2%	5.7%	18.0%	69.1%	11.0%	2.7%	1.3%	0.7%	0.4%	0.3%	0.2%	0.2%	0.1%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%
6	0.3%	0.5%	1.3%	3.9%	12.5%	67.8%	12.0%	3.1%	1.5%	0.9%	0.6%	0.4%	0.3%	0.2%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%
7	0.1%	0.2%	0.5%	1.4%	3.0%	12.6%	65.4%	12.5%	3.3%	1.9%	1.3%	0.9%	0.5%	0.3%	0.2%	0.1%	0.1%	0.1%	0.1%	0.0%
8	0.0%	0.0%	0.2%	0.5%	1.1%	2.7%	13.2%	61.8%	12.2%	4.9%	2.7%	1.7%	1.1%	0.7%	0.4%	0.3%	0.2%	0.1%	0.1%	0.0%
9	0.0%	0.0%	0.1%	0.2%	0.4%	1.0%	2.8%	14.7%	60.3%	22.1%	6.7%	3.5%	2.2%	1.4%	0.9%	0.6%	0.3%	0.2%	0.1%	0.1%
10	0.0%	0.0%	0.0%	0.1%	0.2%	0.5%	1.3%	3.8%	16.5%	56.6%	13.8%	4.8%	3.1%	2.1%	1.5%	0.9%	0.6%	0.3%	0.1%	0.1%
11	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.4%	1.0%	3.0%	7.5%	60.3%	14.3%	5.2%	3.3%	2.3%	1.5%	0.9%	0.5%	0.3%	0.2%
12	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.4%	1.0%	2.7%	7.2%	58.9%	13.9%	5.3%	3.6%	2.5%	1.6%	0.9%	0.4%	0.2%
13	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.5%	1.2%	3.0%	7.2%	56.5%	14.0%	5.6%	3.6%	2.4%	1.5%	0.8%	0.2%
14	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.6%	1.5%	3.2%	7.4%	53.6%	13.6%	5.8%	3.6%	2.2%	1.2%	0.5%
15	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	0.9%	1.8%	3.5%	7.2%	50.5%	13.5%	5.4%	3.1%	1.7%	0.6%
16	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.4%	1.0%	2.0%	3.8%	7.2%	45.7%	12.4%	4.8%	2.5%	0.9%
17	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.5%	1.1%	2.1%	3.8%	7.2%	43.1%	11.7%	3.8%	1.3%
18	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.3%	0.6%	1.1%	2.1%	3.5%	6.5%	38.9%	9.9%	2.0%
19	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	0.6%	1.0%	1.7%	2.9%	5.1%	32.8%	5.5%
20	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.2%	0.3%	1.8%
Above 20	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.2%	0.5%	1.0%	2.0%	3.9%	6.9%	12.7%	19.8%	30.3%	45.8%	86.5%
Executed	2.0%	0.8%	0.6%	0.4%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%

Table 5.2 Short duration orders sequences

Panel I: *DAX15* components

Sequential orders	Stocks	Days	Sequences		Side		Orders	
			Count	%	Bid	Ask	Count	%
Single*							1 864 758	11.3%
2 to 9	15	61	1 884 682	84.8%	50%	50%	7 125 580	43.4%
10 to 49	15	61	320 420	14.4%	50%	50%	5 481 410	33.4%
50 to 99	15	61	12 906	0.58%	53%	47%	858 583	5.22%
100 to 499	15	61	5 448	0.25%	53%	47%	961 625	5.85%
>= 500	14	45	205	0.01%	64%	36%	142 389	0.87%

Panel II: MDAX index components

Sequential orders	Stocks	Days	Sequences		Side		Orders	
			Count	%	Bid	Ask	Count	%
Single*							1 988 269	9.82%
2 to 9	50	61	1 711 485	80.5%	50%	50%	6 408 785	31.7%
10 to 49	50	61	367 679	17.3%	50%	50%	6 806 620	33.6%
50 to 99	50	61	31 183	1.47%	45%	55%	2 091 180	10.3%
100 to 499	50	61	16 316	0.77%	40%	60%	2 681 235	13.2%
>= 500	39	58	362	0.02%	53%	47%	272 559	1.35%

Panel III: SDAX index components

Sequential orders	Stocks	Days	Sequences		Side		Orders	
			Count	%	Bid	Ask	Count	%
Single*							951 309	6.08%
2 to 9	50	61	924 262	73.0%	51%	49%	2 966 018	19.0%
10 to 49	50	61	285 910	22.6%	50%	50%	5 780 373	36.9%
50 to 99	50	61	36 890	2.91%	46%	54%	2 437 941	15.6%
100 to 499	50	61	18 554	1.47%	36%	64%	2 980 939	19.0%
>= 500	41	61	598	0.05%	52%	48%	532 582	3.40%

Table 5.3 Orders sequence general example

KWS stock on 2013-04-26 – Ask side 100 orders sequence

Order	Times		Duration (ms)	Time since previous order cancellation (ms)	Submission Level		Cancellation Level Number	Price	Qty
	Submission	Cancellation			Number	Context			
1	15:05:01.364	15:05:01.374	9.4		1	New	1	277.05	14
2	15:05:01.384	15:05:01.393	9.3	10.3	1	New	1	277.05	14
3	15:05:01.404	15:05:01.413	9.3	10.2	1	New	1	277.05	14
4	15:05:01.423	15:05:01.432	9.3	10.3	1	New	1	277.05	14
5	15:05:01.443	15:05:01.452	9.3	10.2	1	New	1	277.05	14
6	15:05:01.462	15:05:01.472	9.5	10.2	1	New	1	277.05	14
7	15:05:01.482	15:05:01.492	9.5	10.6	1	New	1	277.05	14
8	15:05:01.502	15:05:01.511	9.3	10.3	1	New	1	277.05	14
9	15:05:01.522	15:05:01.531	9.3	10.2	1	New	1	277.05	14
10	15:05:01.541	15:05:01.551	9.3	10.2	1	New	1	277.05	14
11	15:05:01.561	15:05:01.570	9.4	10.1	1	New	1	277.05	14
12	15:05:01.580	15:05:01.590	9.4	10.3	1	New	1	277.05	14
13	15:05:01.600	15:05:01.609	9.3	10.4	1	New	1	277.05	14
14	15:05:01.620	15:05:01.629	9.3	10.2	1	New	1	277.05	14
15	15:05:01.639	15:05:01.649	9.3	10.2	1	New	1	277.05	14
16	15:05:01.659	15:05:01.668	9.3	10.2	1	New	1	277.05	14
17	15:05:01.678	15:05:01.688	9.3	10.2	1	New	1	277.05	14
18	15:05:01.698	15:05:01.707	9.3	10.2	1	New	1	277.05	14
…	…	…	…	…	…	…	…	…	…
83	15:05:02.972	15:05:02.981	9.3	10.3	1	New	1	277.05	14
84	15:05:02.991	15:05:03.001	9.3	10.2	1	New	1	277.05	14
85	15:05:03.011	15:05:03.020	9.3	10.3	1	New	1	277.05	14
86	15:05:03.030	15:05:03.040	9.4	10.2	1	New	1	277.05	14
87	15:05:03.050	15:05:03.059	9.3	10.2	1	New	1	277.05	14
88	15:05:03.070	15:05:03.079	9.3	10.2	1	New	1	277.05	14
89	15:05:03.089	15:05:03.098	9.3	10.2	1	New	1	277.05	14
90	15:05:03.109	15:05:03.118	9.4	10.2	1	New	1	277.05	14
91	15:05:03.128	15:05:03.138	9.3	10.4	1	New	1	277.05	14
92	15:05:03.148	15:05:03.157	9.3	10.3	1	New	1	277.05	14
93	15:05:03.168	15:05:03.178	10.3	10.5	1	New	1	277.05	14
94	15:05:03.189	15:05:03.198	9.5	10.5	1	New	1	277.05	14
95	15:05:03.208	15:05:03.218	9.3	10.3	1	New	1	277.05	14
96	15:05:03.228	15:05:03.237	9.3	10.2	1	New	1	277.05	14
97	15:05:03.247	15:05:03.257	9.3	10.3	1	New	1	277.05	14
98	15:05:03.267	15:05:03.276	9.3	10.3	1	New	1	277.05	14
99	15:05:03.287	15:05:03.296	9.4	10.4	1	New	1	277.05	14
100	15:05:03.306	15:05:03.316	9.2	10.3	1	New	1	277.05	14

Table 5.4 Algorithmic signatures identified in SDAX stocks short duration orders sequences

Type	Count	S	TD	Book side		Orders count		Sequences durations		Rates (sub./sec)		Ranked physical prices proportions										Continuity			Orders submission price levels						
																									Level 1		Level 2		Level 3		≥ 4
				Bid	Ask	Min	Max	Min	Max	Min	Max	1	2	3	4	5	6	7	8	9	10	NC	ETE	OLP	New	Exist	New	Exist	New	Exist	
1.1	19	5	6	21%	79%	100	1335	0:01.9	4:58.8	1	52	100%										100%			100%						
1.2	6	5	6	50%	50%	122	348	0:01.3	5:53.4	1	97	100%										100%					100%				
1.3	7	6	7	0%	100%	116	243	0:00.3	2:11.9	1	404	100%										100%							100%		
2.1.1	2339	48	61	0%	100%	100	245	0:00.1	0:03.7	33	1382	67%	33%										100%		67%	33%					
2.1.2	15	9	14	0%	100%	101	120	0:00.1	0:01.3	87	1259	67%	33%	3%									100%		67%	33%					
2.1.3	6	6	6	0%	100%	118	121	0:01.3	0:01.7	72	94	67%	33%									3%	100%		67%	33%					
2.1.4	19	11	14	0%	100%	116	121	0:00.1	0:01.5	79	1304	67%	33%										100%	3%	67%	33%					
2.1.5	25	17	20	0%	100%	102	242	0:00.1	0:03.2	57	1186	67%	33%										97%		67%	33%					
2.1.6	27	15	21	0%	100%	101	143	0:00.1	0:02.6	55	1275	67%	33%										97%	3%	67%	33%					
2.1.7	15	11	13	0%	100%	107	245	0:00.1	0:02.8	44	1370	63%	33%										100%		67%	33%					
2.1.8	18	14	17	0%	100%	104	137	0:00.1	0:01.8	70	1355	63%	33%	3%									100%		67%	33%					
2.1.9	10	9	10	0%	100%	119	126	0:00.1	0:01.8	68	1049	63%	33%										97%	3%	67%	33%					
2.1.10	8	7	8	0%	100%	102	124	0:00.1	0:02.1	58	1001	63%	33%	3%									97%	3%	67%	33%					
2.1.11	12	9	10	0%	100%	102	123	0:00.1	0:01.3	87	1331	60%	30%	7%	3%								100%		67%	33%					
2.1.12	6	5	6	0%	100%	110	122	0:00.1	0:01.3	86	1295	60%	27%	13%									100%		67%	33%					
2.1.13	11	9	8	0%	100%	108	153	0:00.1	0:01.4	104	1234	53%	27%	13%	7%								100%		67%	33%					
2.2	512	44	57	0%	100%	100	243	0:00.1	0:02.7	48	1406	67%	33%										100%				67%	33%			
2.3	133	26	43	0%	100%	103	142	0:00.1	0:03.6	35	1374	67%	33%										100%						67%	33%	
2.4	88	33	26	0%	100%	100	244	0:00.1	0:02.2	54	1353	67%	33%										100%		33%		33%	33%			
2.5	8	7	6	0%	100%	120	123	0:00.1	0:00.2	647	1201	67%	33%										100%		33%				33%	33%	
2.6	16	11	8	0%	100%	102	141	0:00.1	0:02.3	52	1312	67%	33%										100%				33%		33%	33%	
2.7	54	23	30	0%	100%	101	126	0:00.1	0:01.7	74	1373	67%	33%										100%			33%	33%	33%			
2.8	7	4	4	0%	100%	110	122	0:00.1	0:01.1	106	1261	67%	33%										100%			33%			33%	33%	
2.9	17	9	11	0%	100%	105	142	0:00.1	0:01.3	93	1305	67%	33%										100%					33%	33%	33%	
2.10	18	4	5	0%	100%	102	141	0:00.1	0:02.1	66	1327	67%	33%										100%						33%		67%
2.11	23	13	16	0%	100%	104	123	0:00.1	0:01.6	74	1398	67%	33%										100%								100%
2.12.1	373	12	60	55%	45%	100	2322	0:00.5	19:41.3	1	256	50%	50%									50%		50%	100%						
2.12.2	28	3	23	39%	61%	101	1574	0:54.7	13:02.8	2	2	50%	50%									50%		50%	97%						3%
2.12.3	7	4	6	14%	86%	106	249	0:57.8	2:15.8	2	2	50%	50%									50%		50%	97%						

Count : Number of sequences observed; *S* : Number of different stocks; *TD* : Number of different trading days; *NC* : Non-contiguous; *ETE*: End-to-end; *OLP*: Overlapping; *≥4* : Orders taking place on new or existing fourth best price level and deeper.

Type	Count	S	TD	Book side		Orders count		Sequences durations		Rates (ordr/sec)		Ranked physical prices proportions										Continuity			Orders submission price levels						
																									Level 1		Level 2		Level 3		≥ 4
				Bid	Ask	Min	Max	Min	Max	Min	Max	1	2	3	4	5	6	7	8	9	10	NC	ETE	OLP	New	Exist	New	Exist	New	Exist	
2.12.4	6	3	6	100%	0%	102	367	0:56.0	3:38.3	2	2	50%	47%									50%		50%	100%						
2.12.5	13	6	11	54%	46%	105	2416	1:01.0	22:53.7	1	2	50%	47%	3%								50%		50%	100%						
2.12.6	12	6	11	50%	50%	102	696	0:51.0	6:15.9	2	2	50%	43%	7%								50%		50%	100%						
2.12.7	11	4	9	55%	45%	105	1362	0:57.2	12:06.7	2	2	50%	37%	13%								50%		50%	100%						
2.12.8	7	4	5	57%	43%	104	311	1:05.3	3:17.9	2	2	50%	33%	17%								50%		50%	100%						
2.12.9	7	4	6	43%	57%	104	2075	1:01.2	17:57.3	2	2	50%	27%	23%								50%		50%	100%						
2.12.10	8	4	7	38%	63%	170	2106	1:45.0	20:54.7	2	2	47%	47%									50%		50%	100%						
2.12.11	12	2	10	50%	50%	129	887	1:05.9	9:14.2	2	2	47%	47%									50%		50%	97%						3%
2.12.12	6	4	6	50%	50%	104	487	1:07.3	5:06.8	2	2	47%	47%	3%								50%		50%	100%						
2.12.13	9	6	6	89%	11%	139	1977	1:19.7	20:39.7	2	2	47%	47%	3%	3%							50%		50%	100%						
2.12.14	13	2	12	38%	62%	112	749	0:58.0	6:02.8	2	2	47%	47%	3%								50%		50%	97%						3%
2.12.15	6	3	5	67%	33%	100	663	0:58.9	5:19.8	2	2	47%	47%	3%								50%	3%	50%	97%						3%
2.12.16	12	3	10	33%	67%	106	619	0:51.4	6:08.7	2	2	47%	47%	3%								50%		50%	93%						7%
2.13	13	8	11	46%	54%	106	649	1:04.9	6:25.8	2	2	50%	50%									50%		50%	50%		50%				
3.1	75	27	25	0%	100%	102	142	0:00.1	0:01.5	79	1329	40%	40%	20%									100%		60%	40%					
3.2	17	7	8	0%	100%	117	122	0:00.1	0:01.4	89	1355	40%	40%	20%									100%				60%	40%			
3.3	8	6	6	0%	100%	121	122	0:00.1	0:00.7	176	1313	40%	40%	20%									100%						60%	40%	
3.4	13	6	10	46%	54%	117	381	0:45.5	2:17.4	2	3	33%	33%	33%								33%	33%	33%	100%						
3.5	9	3	3	11%	89%	104	1305	0:27.7	6:17.4	3	6	33%	33%	33%									100%			33%		33%		33%	
3.6	31	18	20	39%	61%	100	894	0:16.5	2:28.8	6	6	33%	33%	33%									100%		33%	33%				33%	
4.1.1	6	4	4	0%	100%	112	140	0:00.1	0:01.2	100	1193	30%	30%	30%	13%								100%		57%	43%					
4.1.2	6	3	4	0%	100%	115	122	0:00.1	0:00.2	522	1216	30%	30%	27%	13%								100%		57%	43%					
5.1	35	12	13	60%	40%	110	1227	0:00.4	6:07.4	3	447	20%	20%	20%	20%	20%							20%	80%	100%						
6.1.1	14	6	8	86%	14%	103	1039	0:01.1	4:13.4	4	220	17%	17%	17%	17%	17%	17%							100%	83%		17%				
6.1.2	8	6	5	63%	38%	131	449	0:32.7	1:54.2	4	4	17%	17%	17%	17%	17%	17%						3%	97%	83%		17%				
6.2	6	2	5	67%	33%	232	247	0:00.3	0:01.7	144	729	17%	17%	17%	17%	17%	17%							100%			83%		17%		
6.3	18	2	3	78%	22%	169	247	0:01.7	0:03.0	65	123	30%	13%	13%	13%	13%	13%					13%	30%	57%	100%						
7.1	10	6	6	50%	50%	102	527	0:29.2	2:30.3	3	4	13%	13%	13%	13%	13%	13%	13%					13%	87%	100%						
8.1	8	2	2	100%	0%	200	436	0:01.6	0:03.8	108	125	30%	10%	10%	10%	10%	10%	10%	10%			10%	30%	60%	100%						
9.1	13	7	7	46%	54%	145	2752	0:01.5	12:45.8	4	131	10%	10%	10%	10%	10%	10%	10%	10%	10%			10%	90%	100%						
10.1	8	5	5	63%	38%	114	598	0:01.4	2:40.3	4	159	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%		10%	90%	100%						

Table 5.5 Algorithmic signatures examples - General

Type	Stock	Date	Time	Duration	Side	Orders count	Rate	Ranked physical prices proportions										Continuity			Orders submission price levels						
																					Level 1		Level 2		Level 3		≥ 4
								1	2	3	4	5	6	7	8	9	10	NC	ETE	OLP	New	Exist	New	Exist	New	Exist	
1.1	KWS	04-26	15:05:01	0:01.942	Ask	100	51	100%										100%			100%						
1.2	BAF	02-22	12:25:49	0:16.835	Bid	122	7	100%										100%					100%				
1.3	P1Z	03-25	10:32:27	0:00.302	Ask	122	404	100%										100%							100%		
2.1.1	CEV	02-06	12:32:30	0:00.082	Ask	101	1227	67%	33%										100%		66%	34%					
2.1.2	O2C	04-16	9:56:12	0:00.787	Ask	114	145	66%	32%	2%									100%		67%	33%					
2.1.3	JUN3	02-20	17:10:23	0:01.483	Ask	118	80	66%	33%	1%								2%	98%		67%	33%					
2.1.4	EV4	03-07	15:45:19	0:00.129	Ask	121	939	66%	33%	1%									98%	2%	67%	33%					
2.1.5	GMM	02-04	11:22:02	0:01.312	Ask	141	107	66%	33%	1%	1%							1%	98%	1%	67%	33%					
2.1.6	SKB	02-27	10:38:25	0:00.116	Ask	112	965	65%	34%	1%									97%	3%	67%	32%			1%		
2.1.7	MLP	04-15	17:11:37	0:00.510	Ask	122	239	63%	34%	2%	1%								98%	2%	66%	34%					
2.1.8	SFQ	04-09	15:53:04	0:00.107	Ask	123	1148	63%	35%	2%									99%	1%	67%	33%					
2.1.9	IVG	02-01	13:28:18	0:00.138	Ask	125	905	63%	33%	2%	1%	1%	1%						96%	4%	66%	33%		1%			
2.1.10	DEX	02-15	10:25:46	0:01.184	Ask	118	100	64%	34%	2%								1%	97%	2%	68%	32%					
2.1.11	B5A	02-26	12:13:15	0:00.118	Ask	123	1041	60%	30%	7%	3%								99%	1%	67%	33%					
2.1.12	JUN3	04-11	12:54:12	0:00.098	Ask	122	1245	61%	27%	12%									100%		67%	33%					
2.1.13	COM	04-11	10:01:32	0:00.120	Ask	119	989	54%	27%	13%	6%								100%		66%	34%					
2.2	GSC1	04-16	9:31:23	0:00.086	Ask	104	1204	67%	33%										100%				65%	35%			
2.3	CEV	03-22	9:36:04	0:01.223	Ask	141	115	67%	33%										100%						67%	33%	
2.4	HAB	03-21	17:15:34	0:00.100	Ask	120	1201	67%	33%										100%		33%		33%	33%			
2.5	CEV	03-04	10:25:58	0:00.186	Ask	120	647	67%	33%										100%		33%				33%	33%	
2.6	HDD	04-02	14:55:17	0:00.973	Ask	141	145	67%	33%										100%				33%		33%	33%	
2.7	SK1A	03-20	12:24:46	0:00.096	Ask	120	1253	67%	33%										100%			33%	33%	33%			
2.8	TTI	03-13	12:06:58	0:00.880	Ask	120	136	67%	33%										100%			33%			33%	33%	
2.9	SIX2	03-26	15:58:51	0:00.104	Ask	121	1159	67%	33%										100%					33%	33%	34%	
2.10	CWC	03-26	12:27:30	0:00.100	Ask	121	1215	67%	33%										100%						33%		67%
2.11	TTI	04-08	16:54:36	0:00.091	Ask	122	1341	67%	33%										100%								100%
2.12.1	HDD	04-12	11:03:37	1:19.326	Ask	126	1.6	50%	50%									50%		50%	100%						
2.12.2	TTI	04-30	11:52:17	2:23.699	Ask	259	1.8	49%	49%	1%								50%	1%	49%	98%						2%
2.12.3	BAF	04-29	16:32:00	2:05.183	Ask	225	1.8	50%	49%									49%		51%	98%	1%	1%				

Type	Stock	Date	Time	Duration	Side	Orders count	Rate	Ranked physical prices proportions										Continuity			Orders submission price levels						
																					Level 1		Level 2		Level 3		≥ 4
								1	2	3	4	5	6	7	8	9	10	NC	ETE	OLP	New	Exist	New	Exist	New	Exist	
2.12.4	B5A	03-06	12:39:24	3:38.308	Bid	367	1.7	49%	48%	2%	2%							50%		50%	100%						
2.12.5	PRA	04-15	14:06:24	22:53.700	Bid	2416	1.8	50%	47%	3%								50%		50%	100%						
2.12.6	2HR	03-12	13:07:39	4:15.050	Bid	474	1.9	50%	43%	7%								50%		50%	100%						
2.12.7	B5A	02-26	17:04:47	0:57.241	Bid	108	1.9	50%	37%	13%								50%		50%	100%						
2.12.8	HAB	03-05	13:08:40	1:05.282	Bid	104	1.6	50%	35%	15%								50%		50%	100%						
2.12.9	B5A	02-01	12:27:02	1:08.041	Ask	110	1.6	50%	27%	23%								50%		50%	100%						
2.12.10	HAB	03-07	16:32:40	4:18.378	Ask	515	2.0	47%	47%	2%	2%	1%	1%					49%	1%	50%	100%						
2.12.11	TTI	04-26	17:01:48	5:14.473	Ask	549	1.7	48%	48%	1%	1%	1%						49%	1%	50%	97%						3%
2.12.12	BAF	04-25	11:28:29	1:07.321	Bid	104	1.5	47%	47%	2%	1%	1%	1%	1%				50%	1%	50%	100%						
2.12.13	HDD	04-26	15:59:50	2:21.530	Bid	251	1.8	47%	46%	4%	3%							49%	1%	50%	100%						
2.12.14	TTI	04-04	11:08:58	2:21.314	Ask	259	1.8	48%	48%	2%								51%		49%	97%						3%
2.12.15	TTI	02-22	15:15:12	2:31.259	Ask	266	1.8	48%	47%	2%	1%	1%						49%	2%	49%	97%						3%
2.12.16	SFQ	04-04	11:20:48	2:08.989	Bid	269	2.1	47%	47%	2%	1%	1%	1%					50%		50%	94%						6%
2.13	TTK	03-27	15:19:20	1:14.439	Ask	125	1.7	50%	50%									50%		50%	50%		50%				
3.1	PMO	03-18	12:25:48	0:00.080	Ask	102	1278	41%	39%	20%									100%		60%	40%					
3.2	CEV	04-18	15:04:59	0:00.096	Ask	120	1248	40%	40%	20%									100%				60%	40%			
3.3	HAB	04-05	11:29:45	0:00.095	Ask	122	1280	40%	40%	20%									100%						60%	40%	
3.4	HAB	03-28	14:02:33	1:37.932	Ask	219	2.2	33%	33%	33%								33%	33%	33%	100%						
3.5	EVD	03-13	15:10:04	1:47.506	Ask	643	6.0	33%	33%	33%									100%			33%		33%		33%	
3.6	2HR	04-15	14:05:00	1:28.119	Ask	529	6.0	33%	33%	33%									100%		33%	33%				33%	
4.1.1	GSC1	04-05	13:51:47	0:00.094	Ask	112	1193	29%	29%	29%	14%								100%		57%	43%					
4.1.2	O2C	04-25	11:00:04	0:00.100	Ask	122	1216	30%	29%	28%	14%								100%		57%	43%					
5.1	SLT	02-05	17:10:38	1:09.276	Ask	228	3.3	20%	20%	20%	20%	20%							20%	80%	100%						
6.1.1	MLP	03-08	10:53:19	2:02.278	Ask	488	4.0	17%	17%	17%	17%	17%	17%							100%	83%		17%				
6.1.2	DAZ	03-07	14:16:24	0:56.765	Ask	225	4.0	17%	17%	16%	16%	16%	16%						2%	98%	84%		16%				
6.2	AB1	03-04	11:32:24	0:00.345	Bid	243	705	17%	17%	17%	16%	16%	16%							100%			83%		17%		
6.3	DEX	03-21	13:03:48	0:01.834	Ask	207	113	29%	15%	15%	14%	14%	14%					13%	29%	58%	100%						
7.1	COM	04-09	12:59:52	0:41.082	Ask	145	3.5	14%	14%	14%	14%	14%	14%	14%	1%				14%	86%	99%		1%				
8.1	KWS	04-24	9:42:38	0:01.606	B	200	125	30%	11%	10%	10%	10%	10%	10%	10%	1%	1%	10%	30%	60%	99%	1%					
9.1	BIO3	03-14	13:04:13	0:46.733	Ask	170	3.6	11%	11%	11%	11%	11%	11%	11%	11%	11%			11%	89%	100%						
10.1	SLT	03-25	15:34:43	0:30.656	Ask	114	3.7	10%	10%	9%	9%	9%	9%	9%	9%	9%	9%	1%	9%	90%	99%		1%				

Table 5.6 Algorithmic signatures examples – Orders

Example 1.2										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	12:25:48.812	12:25:48.983	171.29			2	New	2	4.165	2
2	12:25:48.984	12:25:49.111	126.94	171.99	0.69	2	New	2	4.165	2
3	12:25:49.111	12:25:49.255	143.57	127.56	0.62	2	New	2	4.165	2
4	12:25:49.255	12:25:49.411	155.23	144.15	0.58	2	New	2	4.165	2
5	12:25:49.411	12:25:49.595	184.06	155.93	0.69	2	New	2	4.165	2
...	...	...	...	...	...	...	...	...	...	...
118	12:26:05.164	12:26:05.271	106.75	104.09	0.93	2	New	2	4.165	2
119	12:26:05.272	12:26:05.411	138.43	107.59	0.83	2	New	2	4.165	2
120	12:26:05.411	12:26:05.543	131.72	139.06	0.63	2	New	2	4.165	2
121	12:26:05.543	12:26:05.645	101.88	132.29	0.57	2	New	2	4.165	2
122	12:26:05.647	12:26:05.753	106.11	103.23	1.35	2	New	2	4.165	2

Example 2.1.1										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	12:32:29.888	12:32:29.889	0.78			1	New	1	16.300	81
2	12:32:29.889	12:32:29.890	0.87	0.78	0.00	1	Exist	1	16.300	334
3	12:32:29.890	12:32:29.891	0.98	0.87	0.00	1	New	1	16.295	81
4	12:32:29.891	12:32:29.892	0.77	0.98	0.00	1	New	1	16.300	81
5	12:32:29.892	12:32:29.892	0.73	0.77	0.00	1	Exist	1	16.300	334
6	12:32:29.892	12:32:29.893	0.79	0.73	0.00	1	New	1	16.295	81
...	...	...	...	...	...	...	...	...	...	...
97	12:32:29.968	12:32:29.968	0.83	0.89	0.00	1	New	1	16.300	81
98	12:32:29.968	12:32:29.969	0.70	0.83	0.00	1	Exist	1	16.300	334
99	12:32:29.969	12:32:29.970	0.74	0.70	0.00	1	New	1	16.295	81
100	12:32:29.970	12:32:29.971	0.77	0.74	0.00	1	New	1	16.300	81
101	12:32:29.971	12:32:29.971	0.58	0.77	0.00	1	Exist	1	16.300	334

Example 2.1.13										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	10:01:31.955	10:01:31.960	5.59			1	Exist	1	8.114	452
2	10:01:31.960	10:01:31.976	15.42	5.59	0.00	1	Exist	1	8.114	48
3	10:01:31.976	10:01:31.977	0.89	15.42	0.00	1	Exist	1	8.114	410
4	10:01:31.977	10:01:31.977	0.80	0.89	0.00	1	New	1	8.113	508
...	...	...	...	...	...	...	...	...	...	...
21	10:01:31.990	10:01:31.991	0.81	0.85	0.00	1	Exist	1	8.114	410
22	10:01:31.991	10:01:31.992	0.90	0.81	0.00	1	New	1	8.113	508
23	10:01:31.992	10:01:31.993	0.81	0.90	0.00	1	New	1	8.114	48
24	10:01:31.993	10:01:31.994	0.80	0.81	0.00	1	New	1	8.119	48
25	10:01:31.994	10:01:31.994	0.77	0.80	0.00	1	Exist	1	8.119	472
26	10:01:31.994	10:01:31.995	0.77	0.77	0.00	1	New	1	8.118	890
...	...	...	...	...	...	...	...	...	...	...
117	10:01:32.073	10:01:32.074	0.81	0.81	0.00	1	New	1	8.119	48
118	10:01:32.074	10:01:32.075	0.90	0.81	0.00	1	Exist	1	8.119	472
119	10:01:32.075	10:01:32.076	0.73	0.90	0.00	1	New	1	8.118	890

Example 2.2										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	9:31:22.522	9:31:22.523	0.99			2	Exist	2	75.48	15
2	9:31:22.523	9:31:22.524	0.84	0.99	0.00	2	Exist	2	75.48	160
3	9:31:22.524	9:31:22.525	0.85	0.84	0.00	2	New	2	75.47	181
4	9:31:22.525	9:31:22.526	0.86	0.85	0.00	2	New	2	75.48	15
5	9:31:22.526	9:31:22.527	0.83	0.86	0.00	2	Exist	2	75.48	160
...	...	...	...	...	...	...	...	...	...	...
102	9:31:22.607	9:31:22.608	0.77	0.82	0.00	2	New	2	75.47	181
103	9:31:22.608	9:31:22.609	0.80	0.77	0.00	2	New	2	75.48	15
104	9:31:22.609	9:31:22.610	0.66	0.80	0.00	2	Exist	2	75.48	160

Example 2.4										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	17:15:34.211	17:15:34.211	0.87			2	New	2	7.069	127
2	17:15:34.211	17:15:34.212	0.77	0.87	0.00	2	Exist	2	7.069	700
3	17:15:34.212	17:15:34.213	0.77	0.77	0.00	1	New	1	7.059	127
4	17:15:34.213	17:15:34.214	0.71	0.77	0.00	2	New	2	7.069	127
5	17:15:34.214	17:15:34.214	0.74	0.71	0.00	2	Exist	2	7.069	700
6	17:15:34.214	17:15:34.215	0.74	0.74	0.00	1	New	1	7.059	127
...	...	...	...	...	...	...	...	...	...	...
115	17:15:34.307	17:15:34.308	0.71	0.74	0.00	2	New	2	7.069	127
116	17:15:34.308	17:15:34.308	0.69	0.71	0.00	2	Exist	2	7.069	700
117	17:15:34.308	17:15:34.309	0.73	0.69	0.00	1	New	1	7.059	127
118	17:15:34.309	17:15:34.310	0.70	0.73	0.00	2	New	2	7.069	127
119	17:15:34.310	17:15:34.311	0.84	0.70	0.00	2	Exist	2	7.069	700
120	17:15:34.311	17:15:34.311	0.50	0.84	0.00	1	New	1	7.059	127

Example 2.12.1										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	11:03:36.990	11:03:37.008	17.73			1	New	2	1.718	3167
2	11:03:36.993	11:03:37.010	17.38	2.92	-14.82	1	New	1	1.717	3028
3	11:03:38.233	11:03:38.250	17.72	1239.92	1222.54	1	New	2	1.718	3167
4	11:03:38.235	11:03:38.253	17.56	2.76	-14.97	1	New	1	1.717	3028
...	...	...	...	...	...	...	...	...	...	...
123	11:04:55.033	11:04:55.051	17.72	1277.74	1260.34	1	New	2	1.718	3167
124	11:04:55.036	11:04:55.053	17.82	2.46	-15.26	1	New	1	1.717	3191
125	11:04:56.313	11:04:56.331	17.83	1277.13	1259.31	1	New	2	1.718	3167
126	11:04:56.316	11:04:56.333	17.62	3.09	-14.73	1	New	1	1.717	3191

Example 2.13										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	15:19:20.149	15:19:20.252	102.98			1	New	1	12.545	266
2	15:19:21.332	15:19:21.439	106.54	1182.63	1079.66	2	New	3	12.555	800
3	15:19:21.333	15:19:21.439	106.44	0.80	-105.74	1	New	1	12.545	266
4	15:19:22.519	15:19:22.618	99.01	1185.73	1079.29	2	New	3	12.555	800
...	...	...	...	...	...	...	...	...	...	...
122	15:20:33.358	15:20:33.507	149.27	1180.16	1079.19	2	New	3	12.555	800
123	15:20:33.359	15:20:33.508	148.92	1.04	-148.24	1	New	1	12.545	259
124	15:20:34.587	15:20:34.687	99.48	1228.26	1079.34	2	New	3	12.555	800
125	15:20:34.588	15:20:34.687	99.34	0.89	-98.59	1	New	1	12.545	259

Example 3.1.1										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	12:25:47.762	12:25:47.763	0.96			1	New	1	3.662	344
2	12:25:47.763	12:25:47.764	0.80	0.96	0.00	1	Exist	1	3.662	903
3	12:25:47.764	12:25:47.765	0.86	0.80	0.00	1	New	1	3.661	344
4	12:25:47.765	12:25:47.766	0.86	0.86	0.00	1	Exist	1	3.661	893
5	12:25:47.766	12:25:47.767	1.28	0.86	0.00	1	New	1	3.66	344
6	12:25:47.767	12:25:47.768	0.88	1.28	0.00	1	New	1	3.662	344
…	…	…	…	…	…	…	…	…	…	…
97	12:25:47.838	12:25:47.839	0.79	0.87	0.00	1	Exist	1	3.662	903
98	12:25:47.839	12:25:47.839	0.83	0.79	0.00	1	New	1	3.661	344
99	12:25:47.839	12:25:47.840	1.03	0.83	0.00	1	Exist	1	3.661	893
100	12:25:47.840	12:25:47.841	0.90	1.03	0.00	1	New	1	3.66	344
101	12:25:47.841	12:25:47.842	0.82	0.90	0.00	1	New	1	3.662	344
102	12:25:47.842	12:25:47.843	0.78	0.82	0.00	1	Exist	1	3.662	903

Example 4.1.1										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	13:51:46.608	13:51:46.609	1.14			1	New	1	74.83	29
2	13:51:46.609	13:51:46.611	1.19	1.14	0.00	1	Exist	1	74.83	81
3	13:51:46.611	13:51:46.611	0.87	1.18	0.00	1	New	1	74.82	29
4	13:51:46.611	13:51:46.612	0.90	0.87	0.00	1	Exist	1	74.82	80
5	13:51:46.612	13:51:46.613	0.83	0.90	0.00	1	New	1	74.81	29
…	…	…	…	…	…	…	…	…	…	…
106	13:51:46.697	13:51:46.698	0.84	0.86	0.00	1	New	1	74.83	29
107	13:51:46.698	13:51:46.699	0.80	0.84	0.00	1	Exist	1	74.83	81
108	13:51:46.699	13:51:46.700	0.82	0.80	0.00	1	New	1	74.82	29
109	13:51:46.700	13:51:46.700	0.82	0.82	0.00	1	Exist	1	74.82	80
110	13:51:46.700	13:51:46.701	0.86	0.82	0.00	1	New	1	74.81	29
111	13:51:46.701	13:51:46.702	0.89	0.86	0.00	1	Exist	1	74.81	79
112	13:51:46.702	13:51:46.703	0.83	0.89	0.00	1	New	1	74.8	29

Example 3.4										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	14:02:33.235	14:02:33.246	11.31			1	New	2	7.118	731
2	14:02:33.236	14:02:33.247	11.07	0.96	-10.34	1	New	1	7.117	1915
3	14:02:33.247	14:02:33.248	0.77	11.07	0.00	1	New	1	7.119	1916
4	14:02:34.595	14:02:34.607	11.76	1348.22	1347.45	1	New	2	7.118	731
5	14:02:34.597	14:02:34.608	11.19	1.41	-10.35	1	New	1	7.117	1915
6	14:02:34.608	14:02:34.609	0.90	11.19	0.00	1	New	1	7.119	1916
…	…	…	…	…	…	…	…	…	…	…
214	14:04:09.795	14:04:09.806	11.10	1347.72	1346.85	1	New	2	7.118	730
215	14:04:09.796	14:04:09.807	11.06	0.81	-10.29	1	New	1	7.117	1915
216	14:04:09.807	14:04:09.808	0.98	11.06	0.00	1	New	1	7.119	1916
217	14:04:11.155	14:04:11.167	11.34	1348.31	1347.33	1	New	2	7.118	730
218	14:04:11.156	14:04:11.167	11.21	0.89	-10.44	1	New	1	7.117	1915
219	14:04:11.167	14:04:11.168	0.92	11.21	0.00	1	New	1	7.119	1916

Example 3.5										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	15:10:04.422	15:10:04.454	31.99			3	Exist	3	27	528
2	15:10:04.454	15:10:04.486	32.01	31.99	0.00	2	Exist	2	26.995	528
3	15:10:04.486	15:10:04.780	294.06	32.01	0.00	1	Exist	1	26.95	528
…	…	…	…	…	…	…	…	…	…	…
640	15:11:50.927	15:11:50.959	32.03	426.09	0.00	3	Exist	3	27	528
641	15:11:50.959	15:11:50.992	33.03	32.03	0.00	2	Exist	2	26.995	528
642	15:11:50.992	15:11:51.928	936.07	33.03	0.00	1	Exist	1	26.95	528
643	15:11:51.928	15:11:51.961	33.20	936.07	0.00	3	Exist	3	27	528

Example 3.6										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	14:04:59.678	14:04:59.714	35.90			3	Exist	2	9.93	387
2	14:04:59.714	14:04:59.748	33.99	35.90	0.00	1	Exist	1	9.928	387
3	14:04:59.748	14:05:00.025	277.12	33.99	0.00	1	New	1	9.927	387
…	…	…	…	…	…	…	…	…	…	…
526	14:06:27.559	14:06:27.608	48.97	681.06	0.00	3	Exist	2	9.93	387
527	14:06:27.608	14:06:27.637	29.01	48.97	0.00	1	Exist	1	9.928	387
528	14:06:27.637	14:06:27.797	160.01	29.01	0.00	1	New	1	9.927	387
529	14:06:27.797	14:06:27.826	28.97	160.01	0.00	3	Exist	2	9.93	387

Example 5.1										
Order	Submission	Cancellation	Duration	TSPOS	TSPOC	Level #	Context	Cancel Level	Price	Qty
1	17:10:38.347	17:10:38.963	616.11			1	New	2	34.395	250
2	17:10:38.348	17:10:38.964	616.02	1.00	-615.11	1	New	2	34.39	223
3	17:10:38.963	17:10:39.447	483.96	615.11	-0.91	1	New	1	34.385	250
4	17:10:39.447	17:10:39.948	501.03	483.96	0.00	1	New	2	34.405	250
5	17:10:39.448	17:10:39.949	500.92	1.24	-499.78	1	New	2	34.4	223
6	17:10:39.948	17:10:40.449	501.00	499.78	-1.14	1	New	2	34.395	250
…	…	…	…	…	…	…	…	…	…	…
222	17:11:45.695	17:11:46.224	529.17	1.09	-528.04	1	New	2	34.39	223
223	17:11:46.223	17:11:46.611	387.87	528.04	-1.14	1	New	1	34.385	250
224	17:11:46.611	17:11:47.225	614.11	387.87	0.00	1	New	2	34.405	250
225	17:11:46.612	17:11:47.226	614.12	0.84	-613.27	1	New	2	34.4	223
226	17:11:47.225	17:11:47.622	397.08	613.27	-0.85	1	New	2	34.395	250
227	17:11:47.226	17:11:47.623	397.10	0.85	-396.23	1	New	2	34.39	223
228	17:11:47.622	17:11:48.144	521.97	396.23	-0.87	1	New	1	34.385	250

Table 5.7 SDAX short duration orders standardized liquidity provided

Stock	<10ms		10 ms to 100 ms		100 ms to 1 sec		Total Short Duration	
	Orders	Liquidity	Orders	Liquidity	Orders	Liquidity	Orders	Liquidity
SFQ	544 013	13.9	166 309	67.8	221 276	2 257	931 598	2 339
INH	357 060	26.5	87 812	180.6	287 941	7 425	732 813	7 632
TTI	166 331	11.6	257 879	152.6	225 538	5 886	649 748	6 050
NOEJ	217 443	14.9	250 801	116.7	171 194	4 093	639 438	4 225
BAF	434 698	10.3	124 774	18.3	74 286	201	633 758	229
O2C	493 303	16.3	48 319	10.2	43 728	156	585 350	183
DEX	470 435	17.3	56 664	16.1	53 542	376	580 641	409
B5A	194 803	4.8	305 034	50.4	53 723	315	553 560	370
HDD	333 786	10.9	114 856	25.6	58 990	217	507 632	254
KWS	377 089	21.4	53 204	27.2	57 850	577	488 143	626
DEZ	315 883	8.6	57 760	15.2	79 697	368	453 340	392
DAZ	379 932	7.2	35 933	8.0	30 245	75	446 110	90
MLP	277 869	6.1	36 359	5.3	115 596	317	429 824	328
BDT	315 324	12.9	44 986	15.5	58 840	555	419 150	584
IVG	293 913	8.1	42 302	8.4	41 495	116	377 710	133
P1Z	309 658	6.9	31 466	7.3	28 142	84	369 266	98
GMM	275 206	9.0	34 296	9.6	36 884	253	346 386	272
2HR	229 015	5.8	50 560	10.2	32 753	128	312 328	144
CWC	231 252	8.7	32 893	9.4	44 724	446	308 869	465
AAD	242 226	8.4	25 242	7.5	34 058	207	301 526	223
GSC1	254 588	15.5	31 659	10.9	10 485	57	296 732	84
COM	222 364	8.5	22 897	8.0	46 809	389	292 070	405
HAB	30 482	0.4	233 717	35.0	24 461	121	288 660	157
EVD	186 205	8.2	39 185	13.8	54 568	519	279 958	541
GLJ	214 127	10.1	24 082	7.5	33 855	299	272 064	317

Conclusion

This thesis is intended to be multidisciplinary in several respects. Indeed, as a first step, we developed algorithms allowing us to reconstruct the order book on the basis of information provided by Xetra, a trading system linked to a real stock market.

Subsequently, we extended the data processing to the identification of order book events, and then finally, to the tracking of the orders themselves. It was only after all these steps had been taken that we were able to really focus on the financial issues, which, ultimately, should be the basis of a thesis carried out as part of a PhD in Finance. In doing so, we hope that we have been able to demonstrate how effective data processing is now important to conduct cutting-edge financial researches. Indeed, we consider that this ability to process our main data source and extract the relevant financial information has greatly helped us to innovate, particularly in Chapters 4 and 5.

In Chapter 4, we were able to extend the set of events analyzed in a considerable way compared to what had been done before. Without minimizing our work, this would not have been possible without the methodology developed in Chapter 2, which have allowed us to identify the limit order books events. Similarly, in Chapter 5, we were able to observe and qualify the actual behavior of the algorithms that seem engaged in quote-stuffing operations. This time, we used the limit orders information gathered using the methodology developed in Chapter 3.

Overall, we identify the difficulty of accessing and exploiting data as one of the main causes of the gap that we have unfortunately observed between the scientific literature and what really seems to be taking place in practice, particularly regarding algorithmic and high frequency trading activities. These activities seem to have been the norm in the markets for many years. However, the most widely recognized studies often only consider them globally, in very general and aggregated contexts. In reality, we have often no idea why an order is actually submitted and canceled by a practitioner.

While banks and other market participants have entire teams dedicated to their treatment, it is very difficult for researchers, often individuals, even the most competent and motivated, to extract the relevant information from the data. Thus, although the methodologies developed in this thesis apply mainly to Xetra for the period of our dataset, we hope that they can, at a minimum, serve as a starting point for other researchers who wish to obtain the same type of information, possibly on other markets and other periods.

Bibliography

References

Almgren, Robert F. (2003). "Optimal execution with nonlinear impact functions and trading-enhanced risk." Applied Mathematical Finance **10**(1): 1-18.

Almgren, Robert F., and Neil, Chriss (2000). "Optimal execution of portfolio transactions." Journal of Risk **3**(2): 5-39.

Bacry, E., S. Delattre, M. Hoffmann and J. F. Muzy (2013). "Modelling microstructure noise with mutually exciting point processes." Quantitative Finance **13**(1): 65-77.

Bertsimas, Dimitris, and Andrew W. Lo (1998). "Optimal control of execution costs." Journal of Financial Markets **1**(1): 1-50.

Bowsher, Clive G. (2007). "Modelling security market events in continuous time: Intensity based, multivariate point process models." Journal of Econometrics **141**(2): 876-912.

Carstensen, Lisbeth (2010). "PhD thesis: Hawkes processes and combinatorial transcriptional" University of Copenhagen.

Cheng, Xue, Marina Di Giacinto and Tai-Ho Wang (2017). "Optimal execution with uncertain order fills in Almgren-Chriss framework." Quantitative Finance **17**(1): 55-69.

Cont, Rama, Sasha Stoikov and Rishi Talreja (2010). "A Stochastic Model for Order Book Dynamics." Operations Research **58**(3): 549-563.

Deutsche Börse (2012), " Xetra Release 13.0 Enhanced Broadcast Solution – Interface Specification Programming Version" November 2012.

Deutsche Börse (2013a), "Information on the German High Frequency Trading Bill" Xetra Circular 030/13 April 2013.

Deutsche Börse (2013b), "HFT Bill: Information on Order/Trade Ratio, Excessive System Usage Fees and flagging of Trade Algorithms" Xetra Circular 034/13 April 2013.

Deutsche Börse (2013c), "German High Frequency Trading Act has become effective" Xetra Circular 045/13 May 2013.

Deutsche Börse (2013d), "Further Information on Implementation of single Elements of the German High Frequency Trading Act (HFT)" Xetra Circular 094/13 September 2013.

Deutsche Börse (2013e), "Amendments to Exchange Rules for the Frankfurt Stock Exchange regarding flagging of algorithms and order-to-trade ratios" Xetra Circular 099/13 September 2013.

Dionne, Georges, and Xiaozhou Zhou (2020). "The dynamics of ex-ante weighted spread: an empirical analysis" Quantitative Finance **20**(4): 593-617.

Dionne, Georges, Maria Pacurar, and Xiaozhou Zhou (2015). "Liquidity-adjusted intraday value-at-risk modeling and risk management: An application to data from Deutsche Börse" Journal of Banking and Finance **59**: 202-219.

Egginton, Jared F., Bonnie F. Van Ness, and Robert A. Van Ness (2016), "Quote Stuffing." Financial Management **45**: 583-608.

Embrechts, Paul, Thomas Liniger and Lu Lin (2011). "Multivariate Hawkes processes: an application to financial data." Journal of Applied Probability **48**(A): 367-378.

Fauth, Alexis, and Ciprian A. Tudor (2012). "Modeling First Line Of An Order Book With Multivariate Marked Point Processes." arXiv:1211.4157.

Guo, Qi, and Anatoliy Swishchuk (2020). "Multivariate general compound hawkes processes and their applications in limit order books." Wilmott **2020(**107): 42–51.

Haferkorn, Martin, and Kai Zimmermann (2015). "The German High-Frequency Trading Act: Implications for Market Quality." SSRN Electronic Journal.

Hawkes, Alan G. (1971). "Spectra of Some Self-Exciting and Mutually Exciting Point Processes." Biometrika **58**(1): 83-90.

Hawkes, Alan G., and David Oakes (1974). "A Cluster Process Representation of a Self-Exciting Process." Journal of Applied Probability **11**(3): 493-503.

Hewlett, Patrick (2006). "Clustering of order arrivals, price impact and trade path optimisation." In Workshop on Financial Modeling with Jump processes, Ecole Polytechnique.

Huang, He (2012). "PhD thesis : Modeling Order Book Dynamics Using Queues And Point Processes." The Florida State University.

Huang, Weibing, Charles-Albert Lehalle and Mathieu Rosenbaum (2015). "Simulating and Analyzing Order Book Data: The Queue-Reactive Model." Journal of the American Statistical Association **110**(509): 107-122.

Large, Jeremy (2007). "Measuring the resiliency of an electronic limit order book." Journal of Financial Markets **10**(1): 1-25.

Lu, Xiaofei, and Frédéric Abergel (2018). "High-dimensional hawkes processes for limit order books: Modelling, empirical analysis and numerical calibration", Quantitative Finance, **18**(2): 249-264.

Moallemi, Ciamac C., and Mehmet Sağlam (2013). "Or forum - the cost of latency in high-frequency trading" Operations Research **61**(5): 1070-1086.

Muni Toke, Ioane, and Nakahiro Yoshida (2017). "Modelling intensities of order flows in a limit order book." Quantitative Finance **17**(5): 683-619.

Obizhaeva, Anna A., and Jiang Wang (2012). "Optimal trading strategy and supply/demand dynamics." Journal of Financial Markets **16**(1): 1.

Ogata, Y. (1981). "On Lewis' simulation method for point processes." IEEE Transactions on Information Theory **27**(1): 23-31.

Rambaldi, Marcello, Emmanuel Bacry, and Fabrizio Lillo (2017). "The role of volume in order book dynamics: A multivariate hawkes process analysis" Quantitative Finance **17**(7): 999-922.

Shek, H.H. (2011). "Modeling high frequency market order dynamics using self-excited point process." SSRN Electronic Journal.

Toke, Ioane Muni (2011). “Market Making” in an Order Book Model and Its Impact on the Spread. Econophysics of Order-driven Markets: Proceedings of Econophys-Kolkata V. F. Abergel, B. K. Chakrabarti, A. Chakraborti and M. Mitra. Milano, Springer Milan**:** 49-64.

Toke, Ioane Muni, and Fabrizio Pomponio (2012). "Modelling trades-through in a limit order book using hawkes processes." Economics - The Open-Access, Open-Assessment E-Journal **6**(22): 1-23.

Vinkovskaya, Ekaterina (2014). "PhD thesis : A point process model for the dynamics of limit order books. " Columbia University.

Yergeau, Gabriel (2017). "PhD thesis : Three essays on high-frequency trading algorithms" HEC Montreal.

Yergeau, Gabriel (2018). "Machine learning and high-frequency algorithms during batch auctions." SSRN Electronic Journal.

Zheng, Ban, François Roueff, and Frédéric Abergel (2014). "Modelling bid and ask prices using constrained hawkes processes: Ergodicity and scaling limit" SIAM Journal on Financial Mathematics **5**: 99-136.

Zhou, Xiaozhou (2015). "PhD thesis : High-Frequency Liquidity, Risk Management and Trading Strategy " HEC Montreal.

Other references

Achab, Massil, Emmanuel Bacry, Jean-Francois Muzy, and Marcello Rambaldi (2017). "Analysis of order book flows using a nonparametric estimation of the branching ratio matrix." arXiv.org:1706.03411.

Bacry, Emmanuel, Iacopo Mastromatteo, and Jean-Francois Muzy (2015). "Hawkes processes in finance," Market Microstructure and Liquidity **1**(1).

Biais, Bruno, Pierre Hillion, and Chester Spatt, (1995). "An empirical analysis of the limit order book and the order flow in the Paris Bourse" Journal of Finance **50**: 1655-1689.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan (2014). "High Frequency Trading and Price Discovery." Review of Financial Studies **27**(8): 2267-2306.

Cartea, Alvaro, Richard Payne, Jose Penalva, and Mikel Tapia (2019). "Ultra-fast activity and intraday market quality" Journal of Banking and Finance **99**: 157-181.

Chordia, Tarun, Gideon Saar, Amit Goyal, and Bruce N. Lehmann (2013). "High-frequency Trading." Journal of Financial Markets **16**(4): 637-645.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyamb (2000). "Commonality in Liquidity" Journal of Financial Economics **56**(2): 3-28.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyamb (2001). Market Liquidity and trading activity Journal of Finance **56**(2): 501-530.

Chung, Kee H., and Albert J. Lee (2016). "High-frequency Trading: Review of the Literature and Regulatory Initiatives around the World." Asia-Pacific Journal of Financial Studies **45**(1): 7-33.

Cont, Rama, Arseniy Kukanov and Sasha Stoikov (2014). "The Price Impact of Order Book Events." Journal of Financial Econometrics **12**(1): 47-88.

Da Fonseca, Jose, and Riadh Zaatour (2017). "Correlation and Lead–Lag Relationships in a Hawkes Microstructure Model." Journal of Futures Markets **37**(3): 260-285.

Da Fonseca, José, and Yannick Malevergne (2021). "A simple microstructure model based on the Cox-BESQ process with application to optimal execution policy." Journal of Economic Dynamics and Control, **128**(3).

Filimonov, Vladimir, and Didier Sornette (2015). "Apparent criticality and calibration issues in the Hawkes self-excited point process model: application to high-frequency financial data," Quantitative Finance **15**(8): 1293-1314.

Guo, Qi, Bruno Remillard, and Anatoliy Swishchuk (2020). "Multivariate General Compound Point Processes in Limit Order Books." Risks **8**(3): 98.

Greene, Jason, and Scott Smart (1999). "Liquidity provision and noise trading: Evidence from the investment dartboard column" The Journal of Finance **54**: 1885-1899.

Jain, Pawan, and Steven J. Jordan (2017), "Cancellation Latency: The Good, the Bad, and the Ugly. " Financial Management **46**: 377-407.

Lee, Kyungsub, and Seo Byoung Ki (2017a). "Marked Hawkes process modeling of price dynamics and volatility estimation." Journal of Empirical Finance **40**(3): 174-200.

Lee, Kyungsub, and Seo Byoung Ki (2017b). "Modeling microstructure price dynamics with symmetric Hawkes and diffusion model using ultra-high-frequency stock data." Journal of Economic Dynamics and Control **79**(3): 154-183.

Lu, Xiaofei, and Frédéric Abergel (2017). "Limit order book modelling with high dimensional Hawkes processes" HAL Working Papers hal-01512430.

Lu, Xiaofei, and Frédéric Abergel (2018). "High dimensional Hawkes processes for limit order books Modelling, empirical analysis and numerical calibration" Quantitative Finance **18**(2): 249-264.

Morariu-Patrichi, Maxime, and Mikko S. Pakkanen (2018a). "Hybrid marked point processes: characterisation, existence and uniqueness" Market Microstructure and Liquidity **4**(03n04).

Morariu-Patrichi, Maxime, and Mikko S. Pakkanen (2018b). "State-dependent Hawkes processes and their application to limit order book modelling." arXiv:1809.08060.

Roll, Richard (1984). "A simple implicit measure of the effective bid-ask spread in an efficient market" Journal of Finance **39**: 1127-1139.

Rosu, Ioanid, (2009). "A dynamic model of the limit order book" Review of Financial Studies **22**: 4601-4641.

Swishchuk, Anatoliy, and Aiden Huffman (2018). "General Compound Hawkes Processes in Limit Order Books." Risks **8**(1): 1-25.

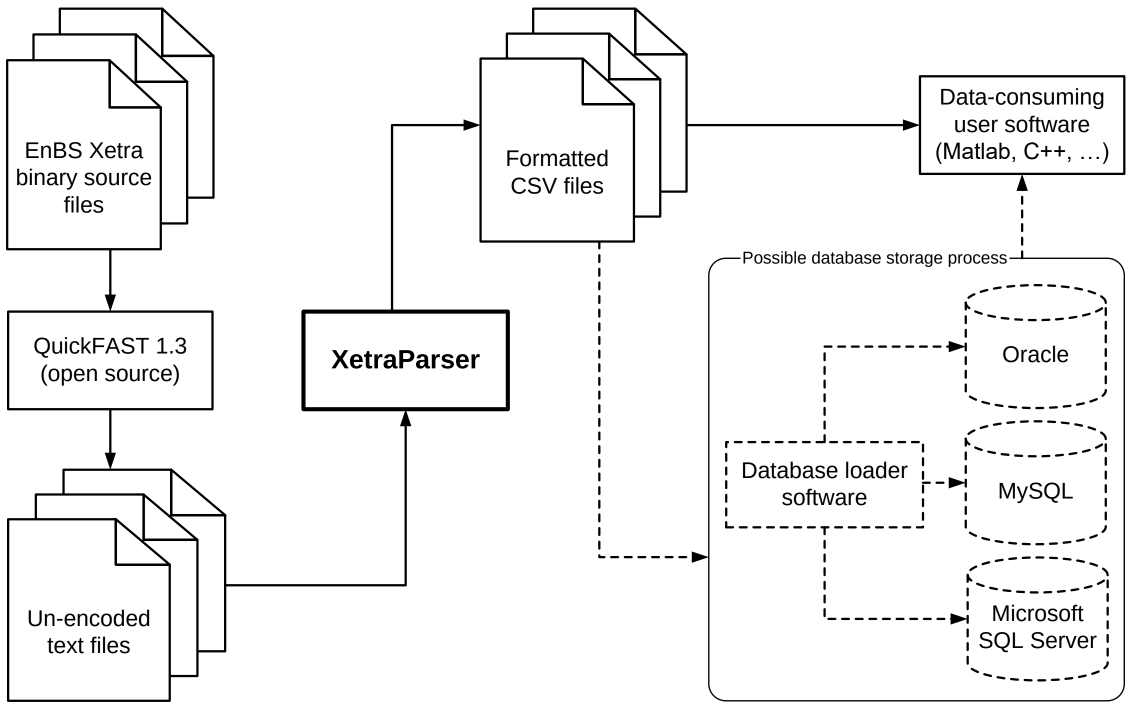
Thibault, Jaisson (2015). "Market impact as anticipation of the order flow imbalance." Quantitative Finance **15**(7): 1123-1135.

Toke, Ioane Muni, and Nakahiro Yoshida (2020). "Marked point processes and intensity ratios for limit order book modeling." arXiv.org:2001.08442.

Appendix 1 : Xetra data processing cycle

XetraParser is the main component of the process visually represented in Figure A1.1.

Figure A1.1 Xetra data processing cycle



Xetra data processing cycle

First, daily Xetra EnBS binary source files are unpacked and decoded using QuickFast, an open source software developed for FIX 5.0 / FAST 1.1 formats processing. Xetra uses these compression / encoding formats for high speed network transmission of real-time market data. This step produces very large un-encoded text files that still can't be used directly. Indeed, under this shape, historical daily data average size is close to 40 GB for millisecond recorded data and 60 GB for more recent microsecond data.

The second step consists in using XetraParser to extract, parse and interpret daily un-encoded text files data. The software can then rebuild the limit order book and retrieve transactions for any stock registered on the Frankfurt Stock Exchange. Since the text files contain a copy of real-time market data, we have to interpret the four main message types originally produced and sent to Deutsche Börse subscribers by Xetra EnBS.

Based on this information, XetraParser uses these messages flows in order to rebuild the order book for any time of a trading day and produces data series allowing us to use it. The output produced by this step consists in formatted CSV files that can finally be used or loaded in a database management system in a third step that will not be described here. It is interesting to note that the size of these output files is quite smaller than the un-encoded source text files (about 6 GB a day for milliseconds reported data).