"Metaorder limit prices in evaluating expected market impact and assessing execution service quality"

| AUTHORS | Viktor Bazylevych (iD) https://orcid.org/0000-0001-8030-8651 |
| | Vitalii Ihnatiuk |

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
23	8	3

BUSINESS PERSPECTIVES

Viktor Bazylevych, Doctor of Science in Economics, Faculty of Economics, Taras Shevchenko National University of Kyiv, Ukraine.

Vitalii Ihnatiuk, Postgraduate student, Faculty of Economics, Taras Shevchenko National University of Kyiv, Ukraine.

Viktor Bazylevych (Ukraine), **Vitalii Ihnatiuk** (Ukraine)

METAORDER LIMIT PRICES IN EVALUATING EXPECTED MARKET IMPACT AND ASSESSING EXECUTION SERVICE QUALITY

Abstract

The paper examines the bias introduced by metaorder limit prices when measuring quality of execution services on financial market. While evaluating the quality of execution services, observed execution costs should be adjusted for metaorder participation rate, size and duration to ensure that they are comparable across execution service providers. One of the exogenous factors which may bias measured execution costs are the different metaorder limit prices in the sample. Currently, there are no proposed methods to normalize for this bias. In the research, the difference in execution costs for metaorders with different limit prices was examined by implementing a limit order book simulation model. It was discovered that the difference in metaorder limit prices is a source of significant heterogeneity in the execution cost distribution. However, we were able to prove that when market agents trade with constant intensities, the difference in execution costs for metaorders with different limit prices is fully explained by their realized participation rate. As a result, financial institution may assess quality of execution services for metaorders without any reservations about differences in metaorders limit prices as long as execution costs are adjusted for different participation rates.

INTRODUCTION

In the modern equity market, financial institutions frequently need to dramatically change the size of their holding of a given stock. However, doing so in one large instantaneous trade may often be too costly or even impossible due to limited supply of market liquidity for an asset. In these cases, market participants resort to slicing large trading orders into smaller pieces and executing them one by one during an extended period of time – hours or days.

These large orders are known in academic literature as metaorders (Farmer et al., 2013). They are executed by a trader – a set of dedicated trading professionals or a system of rules coded up in electronic algorithms. Although these orders are traded during a significant timespan, they still result in changes to supply/demand equilibrium and cause adverse price movement during the execution. This effect is known as market impact (Torre, 1997).

Market impact causes the average execution price for the metaorder to be worse than initially observed on the market. This difference in prices is a markup the financial institution pays for access to liquidity.

According to Collins and Fabozzi's (1991) classification, this markup represents an execution cost portion of variable transaction cost.

Financial institution usually specifies on how a metaorder is to be executed – including limit price and relative speed of trading. Based on these input parameters, the trader decides on the best way to slice a metaorder into pieces and the best time to trade each single slice. Therefore, the market impact and the average execution price that the institution gets for trading the metaorder is subject to the quality of the trader's decisions. This means that the institution should closely monitor the quality of trading services to ensure that it does not pay more than necessary in execution costs.

Still, assessing the performance of traders is plagued by costly, small samples of metaorders and significant noise in the data. Therefore, the following becomes extremely important:

- use of the appropriate metrics for estimating execution costs while using trader services;
- ensuring that the sample size is utilized to its full extent.

The arrival cost metric (also known as 'return proxy') is the measure primarily used in academia and the industry for measuring market impact and variable trading costs (Grinold & Kahn, 1999; Almgren et al., 2005; Bershova & Rakhlin, 2013; Said et al., 2018). It computes the execution price of the order relative to the observed mid-price at the start of the order.

Collins and Fabozzi (1991) claim that the correct measure of transaction costs should satisfy the following criteria:

- it should clearly define an appropriate benchmark for the "fair price" of the transaction;

- the measure of transaction costs should be independent from other factors, which affect execution prices, such as market conditions or order types.

Even though the arrival costs metric is commonly used, it has a drawback with respect to aforementioned requirements: it is affected by a range of market variables and order characteristics.

In order to adjust for these factors, researchers investigated what the 'fair' market impact should be for a given trade. This resulted in a set of well-established formulas to compute expected market impact and expected execution costs (Almgren et al., 2005; Alfonsi et al., 2010; Donier et al., 2015). These formulae use speed of trading, liquidity and volatility of the stock as inputs. Therefore, to adjust for these variables, one could simply normalize observed arrival cost by expected market impact.

For example, when analyzing a trader's metaorder execution, one could reformulate the problem from absolute values $E[arrival\ costs]$ (how much arrival cost this execution had) to relative values $E[arrival\ costs/expected\ costs]$ (how much extra arrival cost this trader caused versus expected).

In the second formulation, arrival costs become adjusted for speed of trading, as the latter is one of the parameters in the expected arrival cost formula:

$$expected\ cost \sim f\left(speed\ of\ trading,\ stock\ liquidity,\ stock\ volatility,\dots\right) \tag{1}$$

Higher arrival cost does not necessarily mean that the trader's execution services are poor – he may have received harder orders to trade than those of his peers (for example, for less liquid stock), which should have a higher execution cost regardless of trader skill. On the other hand, higher adjusted arrival cost is a better indication that the provided services are subpar, as this metric is corrected for order complexity.

However, none of the formulae for expected costs take order type into account. As a result, even when accounting for expected market impact, the adjusted arrival cost metric will be systematically biased if applied to metaorder with different order types (market and limit) and limit prices. A large number of researchers acknowledged this as one of the obvious shortcomings (Cohen et al., 1981; Nevmyvaka et al., 2005).

As a result, when measuring the performance of traders, the institution has to filter out observations of different order types or limit prices and resort to studying only the subsample of executed metaorders. Due to a decreased sample, the financial institution needs to use the services of a potentially underperforming trader longer before being able to determine with statistical confidence that the quality of these services is subpar.

Thus, the aim of this paper is to study how arrival cost should be adjusted so that it can be effectively used when comparing costs of execution for a sample of metaorders with different limit prices.

We attempt to do so by designing a model, which simulates the execution of a large metaorder by a trader. By varying the limit prices of simulated metaorders, we studied the difference in the execution costs caused by this factor and designed a proper normalization procedure.

Our research is structured as follows. In the first section, we review previous research dedicated to metaorders trading and their execution costs. In the second section, we present a set of processes, which attempt to model the execution of a large metaorder by a trader. We examine the differences in the execution costs caused by variability in limit prices and participation rates in the third section. In the fourth section, we provide guidelines on normalization procedure for execution costs of metaorders with different limit prices. We also discuss the limitations of the model and potential areas of future research. Finally, we summarize the findings in the concluding section of the paper.

1. LITERATURE REVIEW

There is an extensive body of scientific research in the differences in execution costs between market and limit orders. For instance, Handa and Shwartz (1991) studied limit order trading and its effect on imbalances and market infrastructure. They found that for traders who trade with lower speeds, the submission of limit orders is preferable to trading using market orders. Biais et al. (1995) have conducted the empirical studies on a set of limit orders sent to the Paris Bourse stock exchange. They observed that large market order trades, which consume outsized liquidity in the limit order book, provide an informational signal to market participants and form expectations on future price movement. Finally, Harris and Hasbrouck (1996) directly compared performance of limit and market orders sent to NYSE. They discovered that on average limit orders perform better than market orders, even though the former are subject to adverse selection risk and higher variability in costs.

However, in these researches, analysis of the execution costs was conducted from the perspective of orders sent directly to the exchange. From metaorder perspective, these exchange orders do not provide the whole picture of metaorder execution cost. In general, a trader is free to slice a metaorder into market or limit exchange orders, as he sees fit as long as the execution price of each slice is within the price discretion provided by the metaorder.

Thus, there is also a set of notable papers dedicated to metaorder market impact. For instance, Almgren et al. (2005) studied functional form of market impact considering metaorder characteristics. His research was continued in the works of Alfonsi et al. (2010), Bershova and Rakhlin (2013), Donier et al. (2015). These studies provide a closed form formula for order expected market impact and for execution cost based on a number of inputs – order characteristics. By normalizing the arrival cost metric by formula-computed expected execution cost, one could account for the differences in these characteristics.

However, none of the formulae of expected costs account for the difference in limit prices. So, one must filter the metaorders with different limit prices from the data set to ensure a homogenous sample when comparing realized arrival costs to the benchmark. For instance, Almgren et al. (2005) pointed out the limit order type bias, but stated that 'analysis of our data set suggests that this effect is not significant' (Almgren et al., 2005, p. 58). This conclusion was made relative to the set of the metaorders being investigated and is unlikely to be extendible to a general case scenario.

Up to this date, the only research conducted into the effect of limits on metaorder market impact was the one of Said et al. (2018). While this research has confirmed some theoretical concepts about market impact for metaorders with different limit prices, it has not identified how to quantify or adjust for this difference.

Our research aims to cross the gap between extensive analysis of the difference in performance between market and limit exchange slices and studies of metaorder market impact. We want to make sure that when computing costs of trading, we could use both limit and market metaorders in the same sample.

Most of the previous studies on market impact were focused on using empirical and historical data. While extremely useful in analyzing costs of individual orders, it is hard to answer the question of what the arrival cost metric would've been if the institution had traded a market instead of limit order. To model these "what-if" scenarios, we need to introduce a set of additional assumptions and directly simulate the underlying order book dynamics.

In doing so, we would model order book dynamics on exchanges, covered in the works of Cont et al. (2010), Farmer et al. (2013), Donier et al. (2015), Abergel et al. (2016). While the most recent models are more closely aligned with the documented empirical facts of metaorder trading, their complex formulaic forms significantly limit flexibility in altering suggested metaorder trading process. We believe that extending the simpler model proposed by Cont et al. (2010) could lead us to understanding the source of the bias in arrival cost metric for limit orders. The latter, in turn, should provide a direction of how the institution could correct for this bias when evaluating trader performance.

2. MODEL

Our model is built based on two interconnected processes.

The first process simulates a limit order book dynamic of a particular asset when demand for liquidity matches the supply. It happens when none of the market participants possesses any convictions about the fair price of the asset. Therefore, price is only influenced by uninformed market trading and resembles random walk.

The second process will model the execution of a buy metaorder following a VWAP/TWAP strategy. It places orders on stock exchange according to the prevailing price and pre-determined strategy schedule. It will cause an increase in demand for liquidity and create an upward pressure on price.

Combining these two processes together, we could directly model effect of buy metaorder execution on liquidity and prices. Note that the side of the metaorder was fixed only for illustrative purposes: the process could easily be extended to a sell case with minor adjustment to equation signs and process interactions.

2.1. Limit order book process

Following the work of Cont et al. (2010), we assumed that at each given moment in time, the stock exchange may be represented as a vector of buy and sell limit exchange orders:

$$x^{buy} = \left(x_{p_1}^{buy}, \ldots, x_{p_k}^{buy}, \ldots, x_{p_N}^{buy}, \ldots \right), \qquad (2)$$

$$x^{sell} = \left(x_{p_1}^{sell}, \ldots, x_{p_k}^{sell}, \ldots, x_{p_N}^{sell}, \ldots \right), \qquad (3)$$

where $x_{p_1}^{buy}$ is the number of limit orders to buy security at price p_1, $x_{p_1}^{sell}$ is the number of limit orders to sell security at price p_1.

Due to the nature of the limit orders, $\forall p_k$: $x_{p_k}^{buy} > 0 \Rightarrow x_{p_k}^{sell} = 0$ (and vice versa), if there are

simultaneous buy and sell orders at the given price, they are crossed and result in a trade.

Limit order books can therefore be expressed more concisely as:

$$x = x^{sell} - x^{buy} = \left(x_{p_1}, \ldots, x_{p_k}, \ldots, x_{p_N} \right), \qquad (4)$$

where $x_{p_k} < 0$ represents $\left| x_{p_k} \right|$ orders to buy and $x_{p_k} > 0$ represents $\left| x_{p_k} \right|$ orders to sell. The current bid price is defined as a maximum price p_B at which x_{p_B} is negative: $p_B = \max \left\{ p_k : x_{p_k} < 0 \right\}$, and the current ask price – minimum price p_A at which x_{p_A} is positive: $p_A = \min \left\{ p_k : x_{p_k} > 0 \right\}$.

The evolution in time of the limit order book follows a continuous-time Markov process with the following set of transitions:

1) arrival of new limit orders with intensity λ:

$$x \to \left(x_{p_1}, \ldots, x_{p_k} + 1, x_{p_N} \right)$$

with rate $\lambda \left(p_k - p_B \right)$ for $p_k > p_B$, (5)

$$x \to \left(x_{p_1}, \ldots, x_{p_k} - 1, x_{p_N} \right)$$

with rate $\lambda \left(p_A - p_k \right)$ for $p_k < p_A$; (6)

2) arrival of new market orders with intensity μ:

$$x \to \left(x_{p_1}, \ldots, x_{p_B} + 1, x_{p_N} \right)$$ with rate μ, (7)

$$x \to \left(x_{p_1}, \ldots, x_{p_A} - 1, x_{p_N} \right)$$ with rate μ; (8)

3) cancellation of orders with intensity dependent on the number of limit orders at the given price:

$$x \to \left(x_{p_1}, \ldots, x_{p_k} - 1, x_{p_N} \right)$$

with rate $\theta \left(p_k - p_B \right) \left| x_{p_k} \right|$ for $p_k > p_B$, (9)

$$x \to \left(x_{p_1}, \ldots, x_{p_k} - 1, x_{p_N} \right)$$

with rate $\theta \left(p_A - p_k \right) \left| x_{p_k} \right|$ for $p_k < p_A$. (10)

In the article, Cont et al. (2010) evaluated the following arrival rates of λ, μ and θ for the Sky Perfect Communication stock as a function of i

– the distance from the opposite quote price:

$$\mu = 0.94, \qquad (11)$$

$$\lambda_i = \frac{k}{i^\alpha}, \text{ where } k = 1.92, \ \alpha = 0.52, \qquad (12)$$

$$\theta(i) = \left\{ 0.71, 0.81, 0.68, 0.56, 0.47 \right\}$$

for $i \le 5$ and $\theta(i) = \theta(5)$ for $i > 5$. (13)

We use the same model and the same parameters to simulate the dynamics of the order book.

Note that, strictly speaking, the process requires maintaining the current book at each single discrete price point on the $(0, \infty]$ interval. However, orders which come outside of the range of possible values for p_B and p_A on the given time horizon have no effect on the evolution of the process. Additionally, the absolute value of the prices (provided they are higher than zero) have no impact on how many price units the stock has moved. Therefore, for simplicity, we modified the price values suggested in the article and assumed a different initial state of order book (depicted in Figure 1).

According to this order book (Figure 1), the quote sizes have an exponentially decaying structure; this is often observed in practice (Abergel et al., 2016). We took special care to ensure that the values of the process never reached beyond the $(0, 200]$ interval in the subsequent simulations. Considering that the model is continuous in its nature, for simulation purposes we discretized the time into transition units of $dt = 0.01$. Since expected number of incoming market orders on one side is equal to $E \left[\# market\ orders \right] = \mu T$ (from properties of Poisson process) and $T = \# transitions \cdot dt$ (due to discretization of time), one market order on average would come in every $1/\mu dt = 106.38$ transitions. This fact allowed us to limit the number of values for which we had to maintain the current list of limit orders and to reduce computational complexity.

Armed with this model description and parameter values, we were able to model the dynamics of the bid and ask prices (Figure 2).
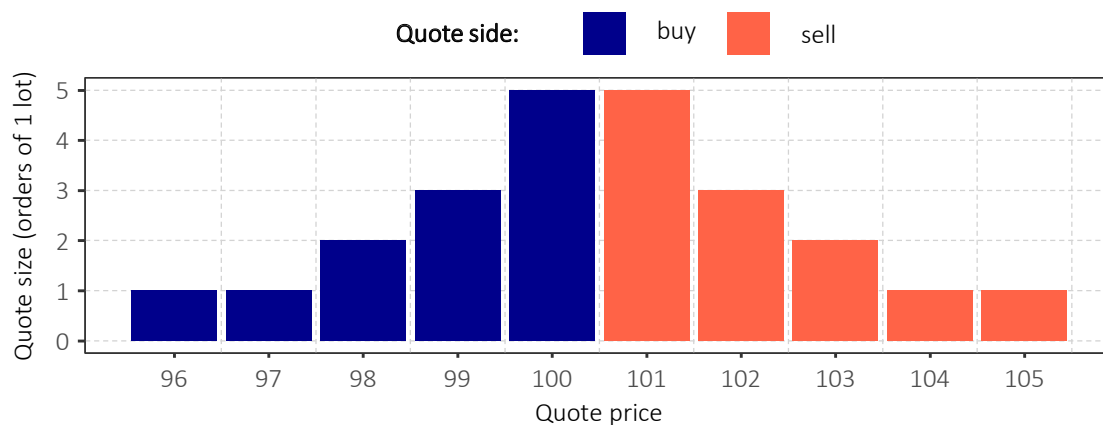
Figure 1. Initial shape of order book
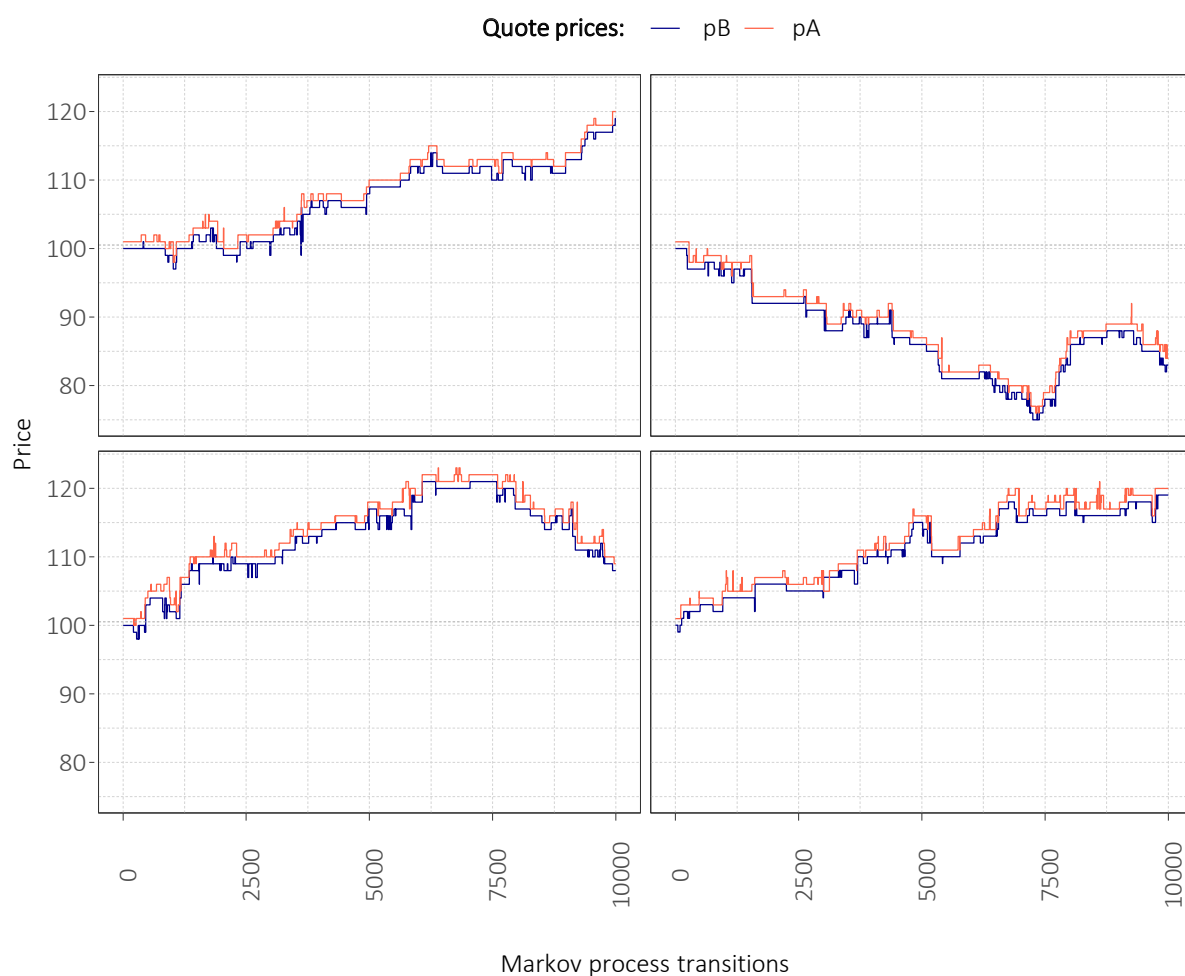
Figure 2. Examples of simulated price dynamics

2.2. Metaorder execution process

Having set up an appropriate model for the limit order books dynamics, we were able to directly model metaorder execution effect on prevailing price. In order to cause market impact, metaorder trading has to represent a short-term shock to the available liquidity and fall outside of what is deemed to be usual trading activity. Therefore, a separate process has to be defined for metaorder execution on top of limit order book transitions.

360

Similar to the previous subsection, we introduced:

$$y = y^{buy} = \left(y_{p_1}, \ldots, y_{p_k}, \ldots, y_{p_N} \right), \tag{14}$$

where y_{p_k} – number of buy slices sent by a trader at price y_{p_k}.

We assumed that the trader attempts to buy stock by trading metaorder with a limit price p_{limit} and following time-weighted-average-price (TWAP) strategy. Since the model does not have any concept of time of day, this strategy behavior should be almost identical to POV or VWAP strategies up to some temporal disturbances caused by the stochastic nature of the process. According to Almgren and Chriss (2001) and Konishi (2002), VWAP execution strategy is the optimal strategy for a risk-neutral trader.

Suppose the trading algorithm wants to buy a slice (one lot) every time window of X time units. For ease of interpretation, we set that time unit = 100 transitions. At the start of that window, our algorithm placed a passive order at the highest passive price determined by the current bid price and limit price of the metaorder:

$$y \rightarrow \left(y_{p_1}, \ldots, y_{p_{passive}} + 1, \ldots, y_{p_N} \right), \tag{15}$$

where $p_{passive} = \min \left\{ p_B, p_{limit} \right\}$ at the moment when passive slice is placed.

Then, the algorithm waited till the end of the X time unit horizon. If the order was traded, the algorithm just remained dormant till the end of the period. If the order didn't result in execution by the end of the period ($y_{p_{passive}}$ was still positive), the algorithm canceled the passive order:

$$y \rightarrow \left(y_{p_1}, \ldots, y_{p_{passive}} - 1, \ldots, y_{p_N} \right) \tag{16}$$

and crossed the spread by aggressively taking liquidity at the ask price (conditional on the ask price being within the limit price) – in essence, sending a market exchange order at price $p_{aggressive}$:

$$x \rightarrow \left(x_{p_1}, \ldots, x_{p_{aggressive}} - \mathbf{1}_{p_{aggressive} \leq p_{limit}}, \ldots, x_{p_N} \right), \tag{17}$$

where $p_{aggressive} = p_a$ at the moment when passive slice is placed.

Taken together, these actions led to exactly 1 slice being traded every X time units, as long as the prevailing prices are within the metaorder limit.

2.3. Combined process

We could express combined process of price limit order dynamics as:

$$z = x - y = x_{p_1} - y_{p_1}, \ldots \tag{18}$$
$$\ldots, x_{p_k} - y_{p_k}, \ldots, x_{p_N} - y_{p_N},$$

where x_{p_k} – difference between number of sell and buy exchange orders at a price p_k not related to execution of metaorder, and y_{p_k} – number of buy slices sent by a metaorder trader at price y_{p_k}.

Dynamics of x_{p_k} will match the process described in the previous subsection. The definition of p_A and p_B will slightly change to reflect the fact that a trader's slices influence market bid and ask prices: $p_B = \max \left\{ p_k : z_{p_k} < 0 \right\}$, $p_A = \min \left\{ p_k : z_{p_k} > 0 \right\}$.

One may note that the limit exchange orders (process z) are no longer homogenuous: a price level may potentially contain both slices of the metaorder under investigation (process y) and exchange orders sent by other market participants (process x). So, if there is an incoming market order to sell, it is no longer appropriate to match it against any order at the bid price. There has to be a rule to identify to which process the contra order belongs.

We overcome this limitation of the model by introducing limit book queues: for each price level, we had a queue of orders with IDs sorted according to their arrival times. Arrival time is defined as the time when an order was placed in the limit book.

Then, we could postulate that:

1) a new limit order is always added to the end of the orders queue of the given price;

2) a new market order is always executed against the oldest order in the queue;

3) in case of cancellation, a random order in the queue is canceled.

Table 1. Description of combined process transitions

Description	Transition	Time component
Process	$z = \left(x_{p_1} - y_{p_1}, \ldots, x_{p_k} - y_{p_k}, \ldots, x_{p_N} - y_{p_N} \right)$ $p_B = \max \left\{ p_k : z_{p_k} < 0 \right\}$ $p_A = \min \left\{ p_k : z_{p_k} > 0 \right\}$	–
Every market transition		
Arrival of new limit order	$x \to \left(x_{p_1}, \ldots, x_{p_k} - 1, x_{p_N} \right)$ with rate $\lambda \left(p_A - p_k \right)$ for $p_k < p_A$ $x \to \left(x_{p_1}, \ldots, x_{p_k} - 1, x_{p_N} \right)$ with rate $\lambda \left(p_A - p_k \right)$ for $p_k < p_A$	A new order is added to the end of queue at price p_k
Arrival of new buy market order	$x \to \left(x_{p_1}, \ldots, x_{p_A} - 1, x_{p_N} \right)$ with rate μ	Order is removed from the start of the queue at price p_A
Arrival of new sell market order	With rate μ: $y \to \left(y_{p_1}, \ldots, y_{p_B} - 1, y_{p_N} \right)$ if the first order in queue at price p_B is metaorder slice; $x \to \left(x_{p_1}, \ldots, x_{p_B} + 1, x_{p_N} \right)$ if the first order in queue at price is not metaorder slice	Order or metaorder slice is removed from the start of the queue at price p_B based on which of the two has earlier arrival time
Cancellation of plain limit order	$x \to \left(x_{p_1}, \ldots, x_{p_k} - 1, x_{p_N} \right)$ with rate $\theta \left(p_k - p_B \right) \left| x_{p_k} \right|$ for $p_k > p_B$, $x \to \left(x_{p_1}, \ldots, x_{p_k} - 1, x_{p_N} \right)$ with rate $\theta \left(p_A - p_k \right) \left| x_{p_k} \right|$ for $p_k < p_A$	Order cancelled at random position in queue at price p_k
Every X time units		
Send aggressive metaorder slice if the passive metaorder slice was not executed in previous *X* time unit period and metaorder limit price is not restrictive	If $y_{p_{passive}} > 0$ and $p_{aggressive} \leq p_{limit}$: $p_{aggressive} = p_a$, $x \to \left(x_{p_1}, \ldots, x_{p_{aggressive}} - 1, \ldots, x_{p_N} \right)$	Metaorder slice is executed against order at the start of the queue at the price $p_{aggressive}$
Cancel the passive metaorder slice if it was not yet executed in previous *X* time unit periods	If $y_{p_{passive}} > 0$: $y \to \left(y_{p_1}, \ldots, y_{p_{passive}} - 1, \ldots, y_{p_N} \right)$	Metaorder slice cancelled from queue at price $p_{passive}$
Send a new passive order	$p_{passive} = \min \left\{ p_B, p_{limit} \right\}$ $y \to \left(y_{p_1}, \ldots, y_{p_{passive}} + 1, \ldots, y_{p_N} \right)$	A new metaorder slice is added to the end of the queue at price $p_{passive}$

These rules represent a plain principle of price-time priority used by most of the exchanges.

The complete set of transitions is presented in Table 1. Figure 3 presents several buy order examples of simulated trading with duration of 100 time units (10,000 market transitions) for an algorithm trading every $X = 10$ time units (every 1,000 transitions).
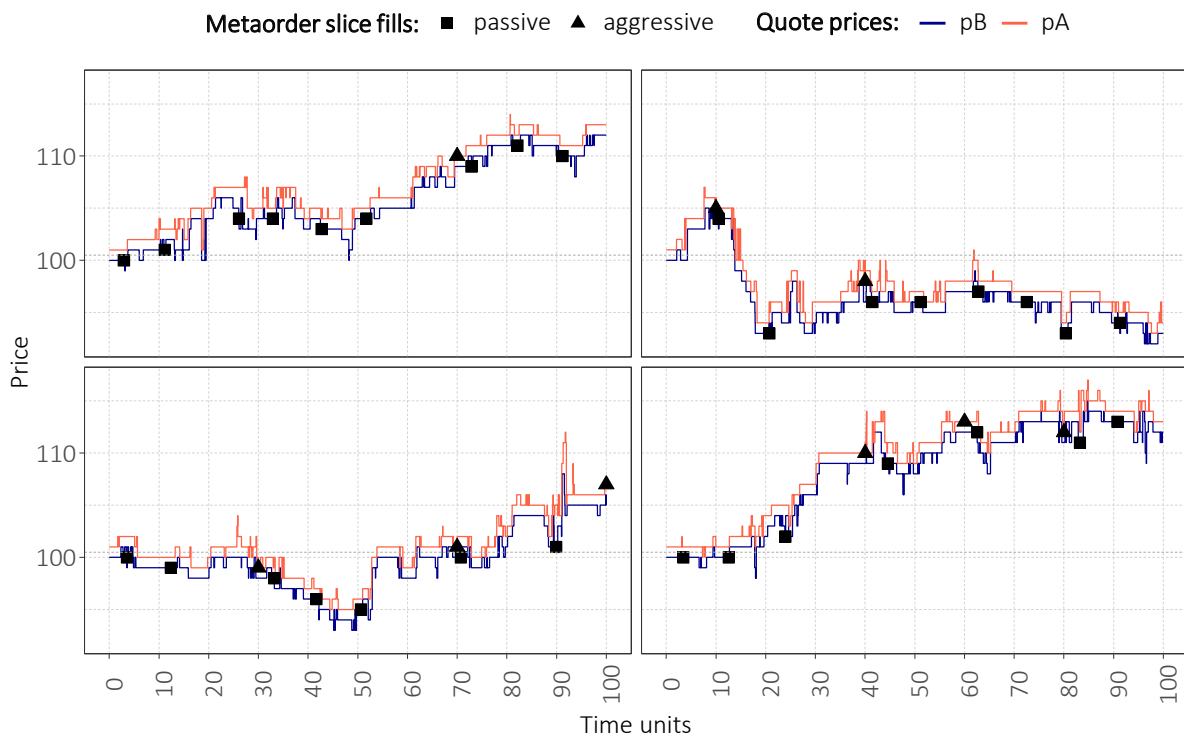
Figure 3. Examples of simulated price dynamics with metaorder traded every 10 time units

3. RESULTS

After we had defined the process, which incorporates market price dynamics and metaorder execution, we were able to investigate how this model behaves under different input parameters.

The two natural variables of interest are the limit order price and speed of trading, expressed as participation rate (Almgren et al., 2005; Toth et al., 2011):

$$Participation\ rate = \frac{Q}{V},\qquad(19)$$

where Q – metaorder size and V – volume traded on the exchange during the execution time of metaorder.

By studying the effect the participation rate has on the cost of execution, we were able to examine whether the model is consistent with a set of empirically observed facts regarding market impact. The latter was used to judge on the quality of model assumptions.

Metaorders trading simulations provided insights into how market impact varies for metaorders with different limit prices. This was used to design an appropriate adjustment to the cost model to account for heterogeneity in metaorder limit prices.

3.1. Speed of trading effect on the model

By changing the frequency with which each slice was traded (selecting various time intervals X), we can directly control the trader liquidity demand and participation rate. Theoretically, the larger the participation, the more impact metaorder trading should have on the price movement and consequent arrival costs.

In order to verify that this holds for the model, we simulated 1,000 orders trading for 100 time units with different frequency of slice trading (every 10, 5, and 2 time periods) and computed arrival costs (or 'return proxy' as defined in Bershova and Rakhlin (2013) and Said et al. (2018)):

$$Arrival\ cost =$$
$$= Side\ sign \cdot \left(\frac{Average\ price}{Arrival\ price} - 1\right),\qquad(20)$$

Table 2. Descriptive statistics for various frequencies (participation rates) of metaorder execution

Time	Arrival cost				Participation rate		IR		
	Mean	**Std. dev.**	**SEM**	**Median**	**Mean**	**Std. dev.**	**Mean**	**Std. dev.**	**SEM**
$X = 10$	2.22%	5.07%	0.16%	2.29%	5.3%	0.4%	5.30%	9.23%	0.29%
$X = 5$	3.95%	5.38%	0.17%	3.58%	10.4%	0.7%	8.83%	9.49%	0.30%
$X = 2$	10.56%	5.23%	0.17%	10.37%	22.9%	1.4%	21.62%	9.41%	0.30%
$X = 1$	19.54%	6.19%	0.20%	19.52%	36.7%	1.8%	38.94%	10.89%	0.34%

where *average price* – average (shares-weighted) price of the execution, *arrival price* – price at the arrival (the moment client decided to trade), *side sign* – adjustment for the direction of the order (+1 for buy and –1 for the sell orders).

We observe that arrival costs increased with increasing frequency of slice trading (Figure 4, Table 2).

Additionally, we investigated the interval return (which is also known as 'market impact' in Bershova and Rakhlin (2013), Donier et al. (2015) and others):

$$Interval\ return =$$

$$= Side\ sign \cdot \left(\frac{Departure\ price}{Arrival\ price} - 1 \right), \qquad (21)$$

where *Departure price* – midpoint of price at the end of executing client order.

A linear relationship between participation rate and interval return (Figure 5) is observed. In real trading, however, there is evidence that this relationship is better described by the declining marginal effect of participation rate on market impact

and concave curve (Toth et al., 2011; Bershova & Rakhlin, 2013; Taranto et al., 2018). This discrepancy highlights an important limitation of the model: it does not take into account that market participants adjust their trading based on the current market activity.

There is an empirically observed fact that when heavy trading occurs at a dislocated price, market participants will try to profit by providing additional liquidity at temporarily inflated price (Bonart & Gould, 2018). The higher the trader participation rate, the more their execution should dislocate the price and incentivize increasing liquidity supply. This will result in a flatter relationship of cost as a function of participation rate. It could be incorporated in our simulation model by setting market and limit order arrival intensities dependent on the current market activity and recent price dislocation. However, it is difficult to determine an appropriate functional form for these relations based on empirical data. We left this topic for future research and focused on the low participation rate range of trading, where this effect is less pronounced. Consequently, the majority of subsequent simulations were conducted using a speed of trading $X = 5$.
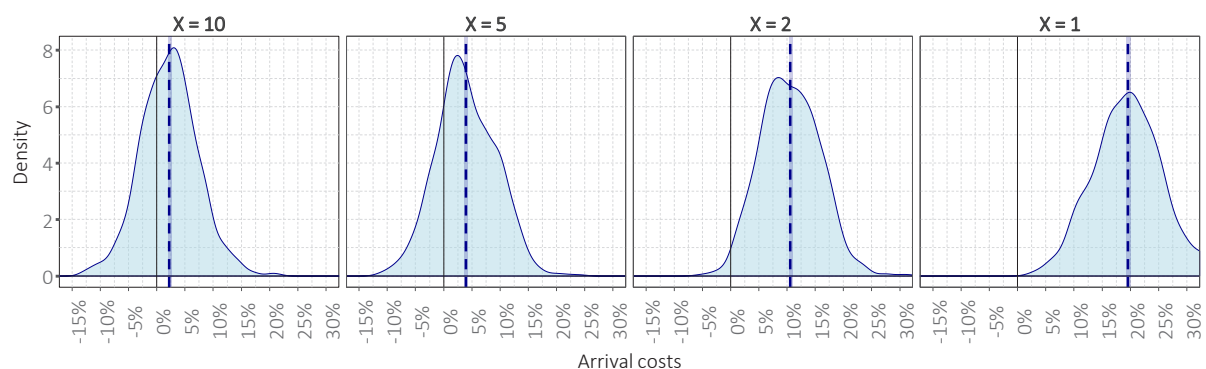
Figure 4. Arrival cost distribution for various frequencies of metaorder trading
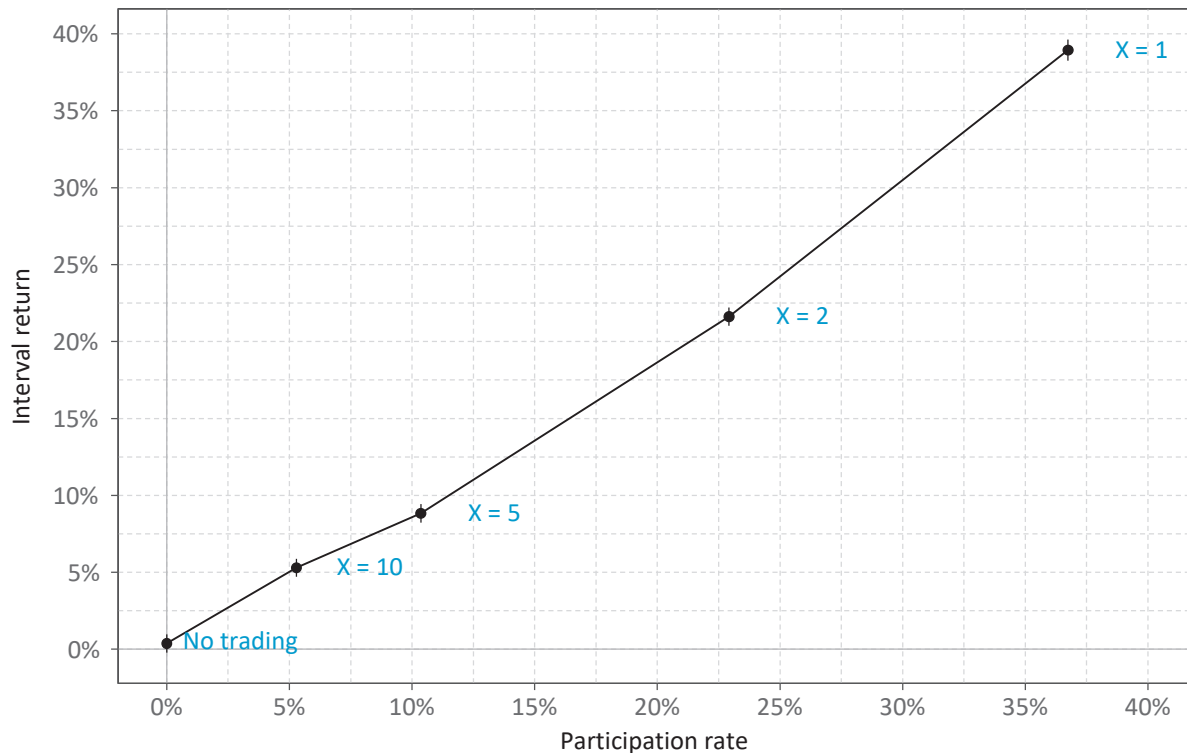
**Figure 5.** Market impact as a function of metaorder participation rate

3.2. Limit price effect

Once we determined that the provided order book process is appropriate for modeling trading impact (with some reservations), we investigated how metaorder limit price affects the distribution of arrival costs and speed of trading. Note that limit price is directly incorporated into the metaorder trading process.

We set the metaorder limit prices to 20, 10, 5 and 0 ticks higher than arrival bid (which translates to limit prices of 19.9%, 9.95%, 4.98% and 0% higher

than arrival price, respectively), then recomputed the distribution of execution arrival costs in each of the cases. Plots of the cost distributions are provided in Figure 6. Limit prices caused the arrival cost distributions to be shifted to the left, resulting in statistically lower arrival costs on average (Table 3). The presence of the limit price cuts off the left tail of the distribution and essentially sets a cap on observed arrival costs.

We operated under the assumption that if a trader is unable to execute an order in the X unit window, he does not speed up his subsequent trading
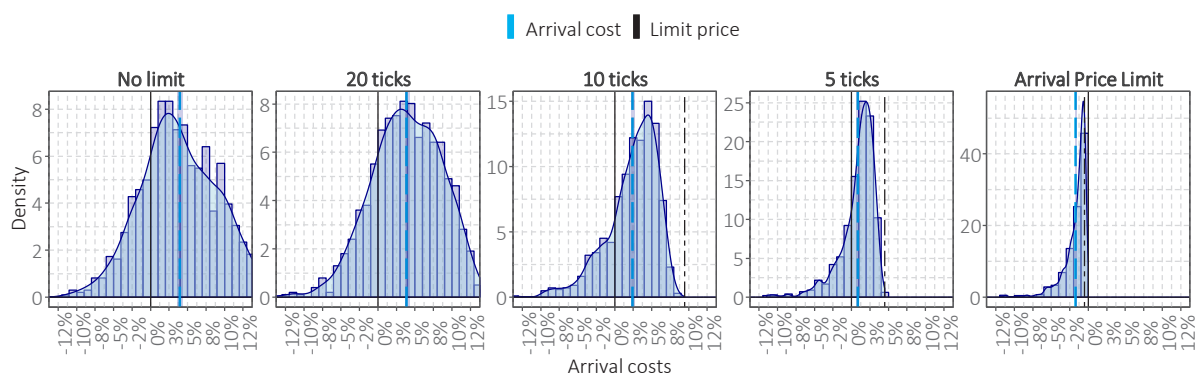
Figure 6. Arrival cost distribution for various limit prices
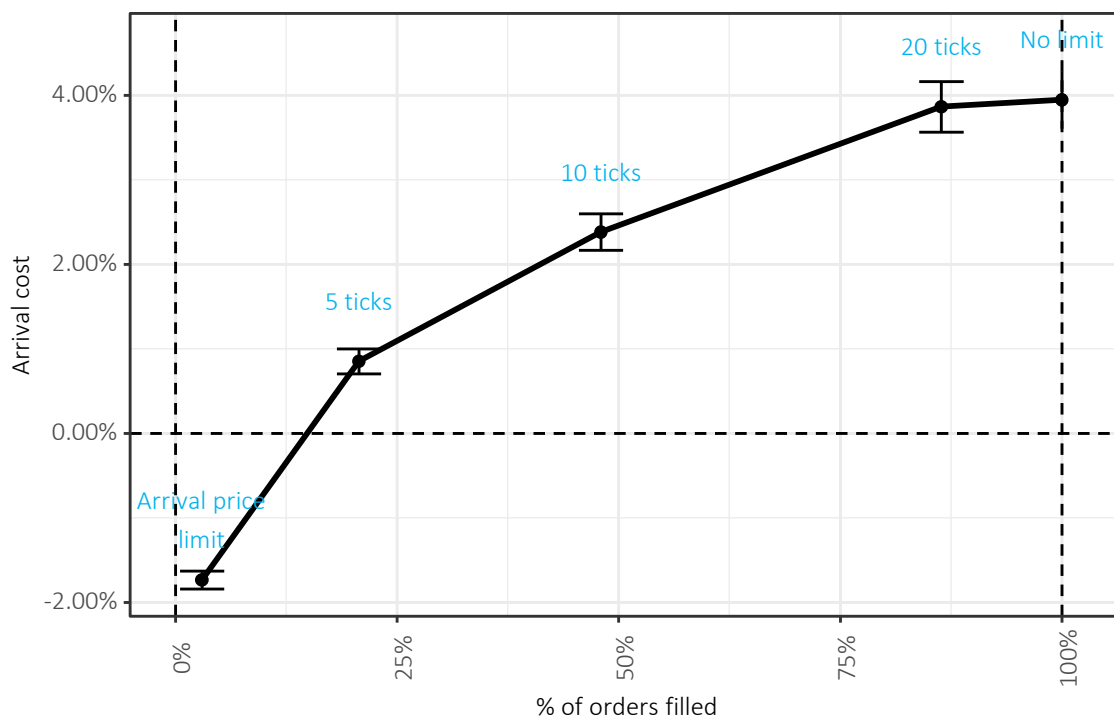
Figure 7. Arrival costs and probability of order to be fully filled

Table 3. Execution cost and trading stats for various metaorder limits

Limit price	Arrival cost			Interval return		Participation rate	% of orders fully filled
	Mean	**Sem**	**t-value**	**Mean**	**Sem**		
No limit	3.95%	0.17%	23.21	8.83%	0.30%	10.36%	100.0%
20 ticks	3.86%	0.15%	25.75	8.85%	0.28%	10.11%	86.4%
10 ticks	2.38%	0.11%	22.09	7.85%	0.26%	8.74%	48.0%
5 ticks	0.85%	0.07%	11.60	6.91%	0.23%	7.01%	20.7%
Arrival price limit	−1.73%	0.05%	−33.04	4.45%	0.23%	4.42%	3.0%

and does not attempt to catch up on the missed volume. As a result, not all orders were fully filled for 20 lots across 100 time units. We observed a tradeoff between the chance of the metaorder to be fully filled and lower costs of execution (Figure 7 and Table 3).

Additionally, the trader exhibited lower participation rate than for no-limit case (Table 3) since he was not able to trade with the intended rate in cases when price became unmarketable. Examination of participation rate versus interval return curve for different limit prices reveals close similarity with the curve derived by changing traders speed of trading (Figure 8). This illustrates that lower

market impact is fully explained by lower average participation rate when trading is restricted by limit price.

We may therefore state that in the absence of second order market effect (such as market participants adjusting their supply and demand of liquidity based on the trader activity), the impact of limit prices on costs is limited to their effect on the trader's speed of trading and does not provide any additional benefits. Therefore, under these assumptions, the trader should be able to achieve the same average costs of trading for market order as for limit order by simply changing his participation rate.
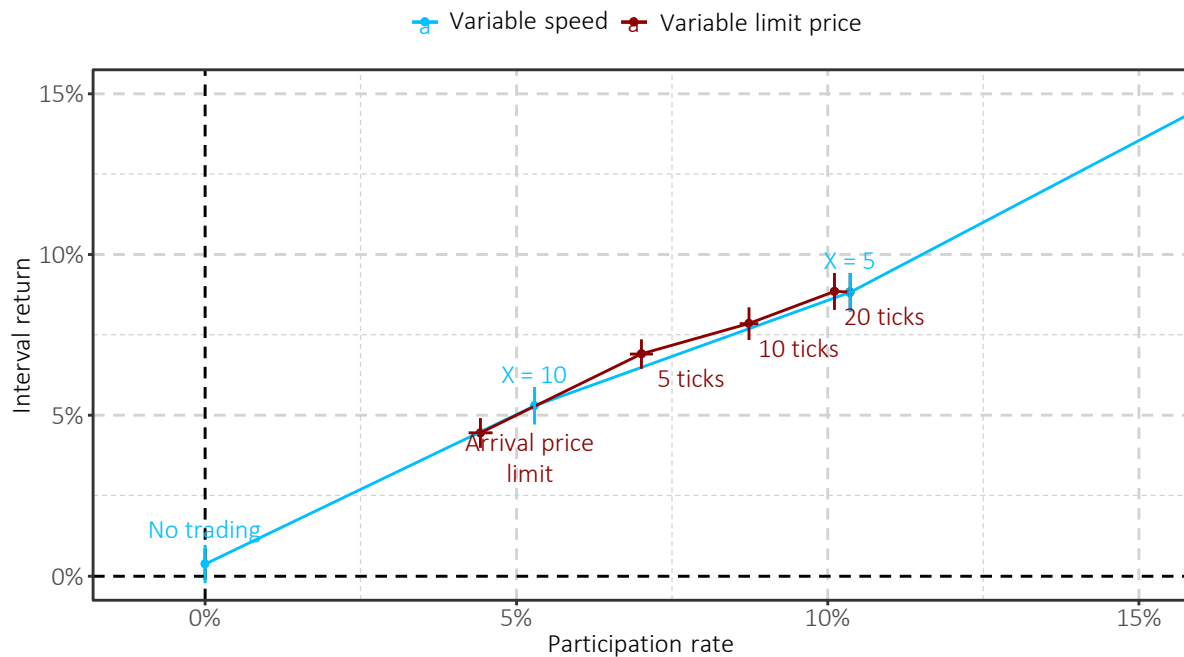
Figure 8. Market impact for various speed of trading and limit prices

4. DISCUSSION

According to the simulations, the effect of the limit price is already incorporated in the metaorder effective participation rate.

Since participation rate is used as a parameter in the market impact cost formulas (Grinold & Kahn, 1999; Almgren et al., 2005), the calculation of expected market impact using the participation rate observed during the life of the order would correctly reflect the presence of the limit prices in metaorders. Therefore, none of the additional adjustments are necessary.

If one wishes to examine the quality of execution of a given algorithm or trader, he could directly compare the observed costs to the expected costs as calculated by models introduced in Grinold and Kahn (1999), Almgren et al. (2005). If the costs are significantly higher than expected costs (i.e. $E\left[arrival\ cost/expected\ cost\right] \gg 1$, one may conclude that the trader provides execution services of low quality.

There should be no additional reservations about metaorder limit prices if other market participants trade with approximately constant intensi-

ties. We believe that making such an assumption during model construction is reasonable, as long as there's no unusual trading activity occurring for a given stock and the metaorder is traded with low participation rate.

However, we acknowledge that this assumption would not hold in a number of cases, for example:

1. If there's significant information-driven trading activity, the assumption regarding constant and symmetrical intensities of market order arrival becomes unrealistic.

2. If the order demands outsized liquidity and represents a significant fraction of market volume, it is unlikely that the market participants would not adjust their behaviour in response to this information.

All of the aforementioned statements could be tested empirically by selecting a universe of metaorders traded with a low speed by a uniform set of traders and then double-binning them by the limit price and realized participation rate. If observed realized costs vary only with respect to realized participation rate dimension and stay constant across the limit price dimension, one could verify

the results of the model on the empirical dataset of metaorders. However, this procedure is extremely data-demanding and it involves a large number of considerations related to cleaning the dataset, selecting the participation rate threshold and ensuring that the selected universe of the metaorders is representative. Solving these problems provides a great area for future studies.

Additionally, in the future, one could attempt to relax the model assumptions by introducing additional processes, which would reflect market agent dynamic reaction to metaorder execution.

There are attempts to design advanced models, which accounts for informational content of the exchange trades (Alfonsi et al., 2010; Farmer et al., 2013; Abergel et al., 2016, etc.), but due to their complex nature, introducing the process of limit metaorder execution to them is a non-trivial task. We believe it is a promising area for future research.

CONCLUSION

We extended the model presented by Cont et al. (2010) to allow for direct modelling of exogenous limit metaorder execution. Upon examining the model behavior, we were able to make the following conclusions.

First, we confirmed that the presence of metaorder limit prices changes the distribution of execution costs and resulting market impact. On average, cost metrics will be lower for metaorders with more restrictive limit prices even after controlling for other initial order characteristics.

This reduction in observed execution costs comes with decreased chance that the order will be fully executed. Investors can build a tradeoff curve between the arrival cost and the chance of the order being fully filled, then decide on the optimal limit placement based on their goals or utility function.

Next, we observed that if market participants do not alter their trading patterns based on the metaorder slices executions, the difference in average arrival costs introduced by limit prices is fully explained by the lower average effective participation rate. Therefore, a trader could achieve the same costs for a market order as for a limit order with any limit price by simply trading slower. Metaorder limit prices do not provide any additional 'optionality' value and serve merely as a matter of convenience and risk bound for a trader.

As a result, it may be possible to account for heterogeneity in execution cost distribution for metaorders with different limit prices by correcting for their effective participation rate. For metaorders traded with low participation rates, investor could simply compare realized execution costs versus expected costs (provided by any market impact model) to assess the quality of execution services without any additional adjustments for the metaorders limit prices.

ACKNOWLEDGEMENTS

REFERENCES

1. Abergel, F., Anane M., Chakraborti A., Jedidi A., & Toke, I. (2016). *Limit Order Books*. Cambridge University Press.

2. Alfonsi, A., Fruth, A., & Schied, A. (2010). Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance, 10*(2), 143-157. https://doi.org/10.1080/14697680802595700

3. Allen, D. (1991). What are transaction costs? *Journal of Behavior and Organization, 14,* 1-18.

4. Almgren, R., & Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk, 3,* 5-39. https://doi.org/10.21314/JOR.2001.041

5. Almgren, R., Thum, C., Hauptmann, E., & Li, H. (2005). Direct estimation of equity market impact. *Risk, 18*(7), 58-62. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.146.1241

6. Berkowitz, S. A., Logue, D. E., & Noser, E. A. (1998). The Total Cost of Transaction on the NYSE. *The Journal of Finance, 43*(1), 97-112. https://doi.org/10.1111/j.1540-6261.1988.tb02591.x

7. Bershova, N., & Rakhlin, D. (2013). The non-linear market impact of large trades: Evidence from buy-side order flow. *Quantitative Finance, 13*(11), 1759-1778. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2197534

8. Biais, B., Hillion, P., & Spatt, C. (1995). An empirical analysis of the limit order book and the order flow in the Paris Bourse. *Journal of Finance, 50*(5), 1655-1689. https://doi.org/10.1111/j.1540-6261.1995.tb05192.x

9. Bonart, J., & Gould, D. (2017). Latency and liquidity provision in a limit order book. *Quantitative Finance, 17*(10), 1601-1616. Retrieved from https://arxiv.org/abs/1511.04116

10. Cohen, K. J., Maier, S. F., & Schwartz, R. A. (1981). Transaction Costs, Order Placement Strategy and Existence of Bid-Ask Spread. *Journal of Political Economy, 89*(2), 287-305. Retrieved from https://www.jstor.org/stable/1833312?seq=1#page_scan_tab_contents

11. Collins, B., & Fabozzi, F. (1991). A Methodology for Measuring Transaction Costs. *Financial Analyst Journal, 47*(2), 27-36. Retrieved from https://www.jstor.org/stable/4479413?seq=1#page_scan_tab_contents

12. Cont, R., Stoikov, S., & Talreja, R. A. (2010). Stochastic Model for Order Book Dynamics. *Operations Research, 58*(3), 549-563. Retrieved from https://www.jstor.org/stable/40792679?seq=1#page_scan_tab_contents

13. Donier, J., Bonart, J., Mastromatteo, I., & Bouchaud, J. (2015). A fully consistent, minimal model for non-linear market impact. *Quantitative Finance, 15*(7), 1109-1121. Retrieved from https://arxiv.org/abs/1412.0141

14. Farmer, J. D., Gerig, A., Lillo, F., & Waelbroeck, H. (2013). How efficiency shapes market impact. *Quantitative Finance, 13*(11), 1743-1758. Retrieved from https://arxiv.org/abs/1102.5457

15. Grinold, R., & Kahn, R. (1999). *Active Portfolio Management* (2nd ed.) (616 p.). New York: McGraw-Hill.

16. Handa, P., & Schwartz, R. (1996). Limit order trading. *Journal of Finance, 51,* 1835-1861.

17. Harris, L., & Hasbrouck J. (1996). Market vs. Limit orders: The SuperDOT evidence on order submission strategy. *Journal of Financial and Quantitative Analysis, 31*(2), 213-231. https://doi.org/10.2307/2331180

18. Konishi, H. (2002). Optimal Slice of a VWAP Trade. *Journal of Financial Markets, 5,* 197-221. https://doi.org/10.1016/S1386-4181(01)00023-4

19. Nevmyvaka, Y., Kearns, M., Papandreou, M., & Sycara, K. (2005). *Electronic trading in order-driven markets: efficient execution*. Paper presented at Seventh IEEE International Conference on E-Commerce Technology (CEC'05) (pp. 190-197). Munich, Germany.

20. Said, E., Ayed, A., Husson, A., & Abergel, F. (2018). *Market Impact: A systematic study of limit orders* (Working Papers hal-01561128). HAL.

21. Taranto, D., Bormetti, G., Bouchaud, J., Lillo, F., & Toth B. (2018). Linear models for the impact of order flow on prices I. History dependent impact models. *Quantitative Finance, 18*(6), 903-915.

22. Torre, N. G. (1997). *Market Impact Model Handbook*. BARRA Inc., Berkeley.

23. Toth, B., Lemperiere, Y., Deremble, C., Lataillade, J., Kockelkoren, J., & Bouchaud, J. (2011). Anomalous price impact and the critical nature of liquidity in financial markets. *Physical Review, X1*(2), 1-11. Retrieved from https://arxiv.org/ct?url=https%3A%2F%2Fdx.doi.org%2F10.1103%2FPhysRevX.1.021006&v=b6b9b28a