HAL
open science

Market Impact: A systematic study of limit orders

Emilio Said, Ahmed Bel Hadj Ayed, Alexandre Husson, Frederic Abergel

To cite this version:

Emilio Said, Ahmed Bel Hadj Ayed, Alexandre Husson, Frederic Abergel. Market Impact: A systematic study of limit orders. 2017. hal-01561128v1

HAL Id: hal-01561128

https://hal.science/hal-01561128v1

Preprint submitted on 12 Jul 2017 (v1), last revised 14 May 2022 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Market Impact: A systematic study of limit orders

Emilio Said[*,**], Ahmed Bel Hadj Ayed[*], Alexandre Husson[*], and Frederic Abergel[**]

[*]Global Markets, BNP Paribas, Paris
[**]Chair of Quantitative Finance, MICS laboratory, CentraleSupelec, Paris

July 12, 2017

Abstract

This paper is devoted to the important yet little explored subject of the market impact of *limit orders*. Our analysis is based on a proprietary database of metaorders, that is, large orders that are split into smaller pieces. It is divided into two distinct parts: the first addresses the case of aggressive limit orders and the second, that of passive limit orders. For both order types, we confirm a square root temporary impact in the participation rate. In the case of aggressive orders, we also study the price relaxation after the end of the metaorder, and show that the long-term impact reaches a level of approximately two-thirds of the maximum impact. Finally, we propose a functional form for market impact given by $\sigma\sqrt{\frac{Q}{V}}$, where σ and $\frac{Q}{V}$ respectively stand for the volatility and the participation rate.

1 Introduction

It is a commonly acknowledged fact that market prices move because of the execution of a trade - they go up for a large buy order and decrease for a large sell order. This is, losely stated, the phenomenon known as *market impact*.

The market impact of large trading orders that are split into pieces - better known as **metaorders** - and executed incrementally by a succession of orders of smaller sizes is crucial in describing and controlling the behaviour of modern financial markets. Being able to quantify this impact is clearly a question of great relevance in studying the price formation process, and it has also become a major practical issue for optimal trading. Indeed, to know whether a trade will be profitable, it is essential to monitor transaction costs, which are directly linked to market impact. Measuring and modelling market impact has therefore become a central question of interest, both for researchers and practitioners, in the field of market microstructure.

Compared to the importance of the subject, there only exists a rather scarce literature pertaining to the empirical estimation of market impact, mostly due to poor data availability. In fact, trades and quotes, or even order book data are not sufficient to perform the analysis: a clear identification and time stamping allowing the reconstruction of metaorders is required. Metaorders have started being recorded as such in a systematic way only recently, and mostly in proprietary databases that are not readily accessible to academic researchers in the field of market microstructure. The analyses presented in [Almgren et al., 2005], [Moro et al., 2009], [Gatheral, 2010], [Toth et al., 2011], [Bershova and Rakhlin, 2013], [Bacry et al., 2015], [Gomes and Waelbroeck, 2015] essentially cover all that is published about the market impact of large orders.

Although difficult to measure in practice, market impact has been studied from a theoretical point of view. In an economic theory perspective, the information held by investors, which governs theirs decisions, should have some predictive power over future prices. This point was thoroughly investigated by [Farmer et al., 2013], a paper which we find enlightening and use as reference for the theoretical measurements of market impact. In a related study, [Bershova and Rakhlin, 2013] uses their own proprietary database to perform an empirical analysis of a set of large institutional orders, and validates some predictions of the [Farmer et al., 2013] model. Such a comparison will also be performed in the present work.

Our paper is a contribution to this strand of research, with a specific focus on the market impact of *limit orders*. As a matter of fact, market orders are generally not used by institutional investors because of the lack of control they imply. On the contrary, limit orders, whether they are *aggressive* - crossing the spread - or *passive*, form the vast majority of orders actually sent to the market during the execution of a large trade. As such, they should be the main subject of interest in a study of market impact. To the best of our knowledge, this paper is the first academic study of market impact with such an emphasis.

The statistics we present are made on a proprietary database, formed of appropriately selected limit orders executed on the European equity market between January and August 2016, at an intraday scale. The originality of our approach lies in the fact that the study is not carried out directly on raw order data, but rather relies on an algorithmic identification and reconstruction of metaorders from the database of all orders.

The paper is organized as follows: Section 2 recalls for further reference the framework and main results of the theoretical, agent-based market impact model developed in [Farmer et al., 2013]. Section 3 introduces our main definitions of metaorders and market impact measures. Section 4 presents our empirical results: They confirm that aggressive limit orders behave in agreement with some *stylized* facts already established in the literature, and shed a new light on the behaviour of passive orders. Section 5 is a discussion of our results and their implications, including some comparisons with the literature.

2 The Market Impact Model of [Farmer et al., 2013]

In this section, the main results of [Farmer et al., 2013] are recalled. This model presents a theory for the market impact of large trading orders that are split into pieces and executed incrementally, better known as **metaorders**. The central goal of the model is to understand how order splitting affects market impact.

2.1 Model description

- A filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is given.
- At $t = 0$, before the opening of the market, the $\mathcal{K}$ long-term traders have a common information signal α and each trader $k = 1, ..., \mathcal{K}$ formulate an order of size $0 \leq n_k(\alpha) \leq \frac{M}{\mathcal{K}}$ and submit them to the algorithmic trading firm that bundles them together into a metaorder of size $N = \sum_{k=1}^{\mathcal{K}} n_k$, where $M, \mathcal{K}, \frac{M}{\mathcal{K}}$ are all large positive numbers.
- With probability μ the signal α has nonzero support over a continuous interval $-\alpha_{max} \leq \alpha \leq \alpha_{max}$, where $0 < \alpha_{max} < \infty$. With probability $1 - \mu$, $\alpha = 0$ which implies that there will be no metaorder, i.e. $n_k = 0$ for all k.

- There are N auctions following each other $t = 1, ..., N$, $N \leq M$, N representing the number of orders necessary to execute fully the metaorder.
- At each auction $t = 1, ..., N$, each day trader submits a market order of size $F(\eta_t)$ where (η_t) is a zero mean iid random process and F an increasing function whose functional form is not important. The market makers observe the order flow and formulate quotes independently.
- At $t = N + 1$, the last instant in the game, i.e. the instant after the metaorder is fully executed, is announced with the final price $\widetilde{X_N}$:

$$\widetilde{X_N} = G(X_0, \alpha, \eta_1, ..., \eta_N), \text{ if } N > 0$$

  where X_0 is the initial price and G a function whose form is not important. **We will use the tilde notation to refer to a quantity depending on** (η_t).
- **Long-term traders** only have the information signal α, they know that others traders have received the same signal and submit an order of size $n_k(\alpha)$ independently.
- **Day traders** receive a private information η_t at each period and mechanically submit an order of size $F(\eta_t)$, they play the role of *noise traders*.
- **Market makers** are competitive liquidity providers, able to take past order flow and prices into account in setting their quotes. They know the initial price X_0, the information distributions $P(\alpha)$ and $\widehat{P}(\eta_t)$, the probability μ that a metaorder is present and the function F relating the day trader's information to their order size. During the course of the game they can update their prior μ to make a time dependant estimate μ'_t.

The respective roles of each agent are summarized in the table below (extracted from the article).

Agent	action	period	information
long-term investor	submits order $n_k(\alpha)$ to ATF	0	α
algorithmic trading firm (ATF)	submits lot from metaorder	t	$N = \sum_{k=1}^{\mathcal{K}} n_k(\alpha)$
day trader	submits order F(η_t)	t	η_t
market maker (MM)	price quote, MM with best quote, executes transactions at price $\widetilde{S}_t$	t	combined order $sgn(\alpha) + F(\eta_k)$

Table 1: *Agents in the model.*
(extracted from [Farmer et al., 2013])

The long-term investor is the key agent, who must choose an order size n_k at the beginning of the game based on information α. The market makers are competitive profit optimizers who set prices (but do not know α). The day trader and the algorithmic trading firm are mechanical agents, the day trader provides a background of noisy order flow and the algorithmic trading firm bundles up the long-term trades and submits them in lots of equal size.

2.2 Notations

- As regards statistical averages, $\widetilde{S}_t$ denotes a specific realization of transaction prices, whereas S_t stands for an average price over the day trader's signal η_i. Likewise, $\widetilde{X_N}$ is a specific realization of the final price at the end of the metaorder whereas X_N is an average price over the day trader's signal η_i;

- The final price averaging over η_t is denoted by X_N i.e $X_N = \mathbb{E}[\widetilde{X_N}]$;
- $X_{t-1} = S_{t-1} - R^-_{t-1}, t = 2, ..., M$ and $X_M = S_M$ thus $R^-_M = 0$;
- $S_t = X_0 + \sum_{i=0}^{t-1} R^+_i, t = 1, ..., M$, where $R^+_t = S_{t+1} - S_t$ and $R^-_t = S_t - X_t$ are the corresponding incremental average impacts;
- $\mathcal{I}_t = S_t - X_0$ is the **average immediate impact** at t and $I_N = X_N - X_0$ denotes the **average permanent impact**;
- Let $m \in \{0, 1\}$ an indicator variable where $m = 1$ if the metaorder is present and $m = 0$ if it is absent
- We will consider $p_t = \mathbb{P}(t \leq N < t + 1 \,|\, m = 1)$ at each period t and $\mathcal{P}_t$ the probability that the metaorder will continue given that it is still active at timestep t, i.e.

$$\mathcal{P}_t = \frac{\sum_{i \geq t+1} p_i}{\sum_{i \geq t} p_i} = \mathbb{P}(N \geq t + 1 | N \geq t, m = 1)$$

.

The possible price paths for a buy are summarized in the graph below:

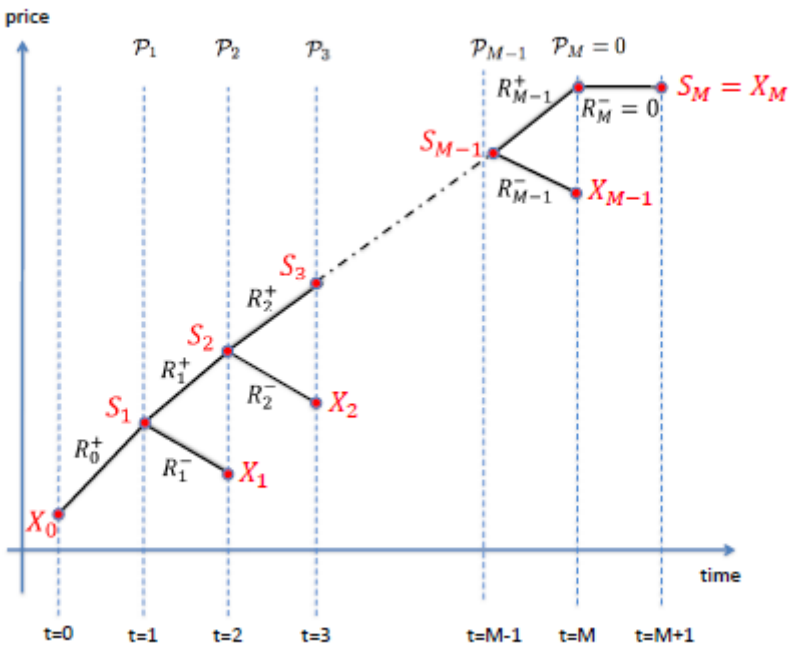


Figure 1: *Tree of possible price paths for a buy metaorder for different sizes N.* (extracted from [Farmer et al., 2013])

In Figure 1, the metaorder is supposed to be present and only expected price paths, averaged over the day trader's information (which is why the notation does not include tildes), are shown.

The price is initially X_0, after the first lot is executed it is $S_1 = X_0 + R_0^+$. If $N = 1$ it is finished and the price reverts to $X_1 = S_1 - R_1^-$, but if $N > 1$ another lot is executed and it rises to $S_2 = S_1 + R_1^+$. This proceeds similarly until the execution of the metaorder is completed. At any given point the probability that the metaorder has size $N > t$, i.e. that the order continues, is $\mathcal{P}_t$. A typical price path (rather than the expected price paths shown here) subject to a large day trader's noisy information signal would look more like a random walk with a time-varying drift caused by the impact of the metaorder.

2.3 Martingale condition

In this section, $(\widetilde{S}_t)$ is supposed to be a martingale: $\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t] = \widetilde{S}_t$. Therefore, the following properties hold true:

- $\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t] = \mathbb{P}(m = 1|\mathcal{F}_t)\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t, m = 1] + \mathbb{P}(m = 0|\mathcal{F}_t)\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t, m = 0]$, and:

$$\mathbb{P}(m = 1|\mathcal{F}_t) = \mu'_t.$$

  due to the fact that μ'_t is the market maker's best estimate of the probability that the metaorder is present;

- For the price in the next period, it is necessary to average over three possibilities :

$$\begin{aligned}\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t, m = 1] &= \mathbb{P}(N = t|\mathcal{F}_t, m = 1)\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t, m = 1, N = t] \\ &+ \mathbb{P}(N > t|\mathcal{F}_t, m = 1)\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t, m = 1, N > t]\end{aligned}$$

  with :

  - ⋆ $\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t, m = 1, N > t] = \widetilde{S}_t + R_t^+$
  - ⋆ $\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t, m = 1, N = t] = \widetilde{S}_t - R_t^-$
  - ⋆ $\mathbb{E}[\widetilde{S}_{t+1}|\mathcal{F}_t, m = 0] = \widetilde{S}_t$
  - ⋆ $\mathbb{P}(N > t|\mathcal{F}_t, m = 1) = \mathbb{P}(N \geq t + 1|N \geq t, m = 1) = \mathcal{P}_t$.

  thus :

$$\begin{gathered}\widetilde{S}_t = \mu'_t(\mathcal{P}_t(\widetilde{S}_t + R_t^+) + (1 - \mathcal{P}_t)(\widetilde{S}_t - R_t^-)) + (1 - \mu'_t)\widetilde{S}_t \\ \mathcal{P}_t R_t^+ - (1 - \mathcal{P}_t)R_t^- = 0\end{gathered}$$

  Since $(\widetilde{S}_t)$ is a martingale, there holds :

Proposition 1.

$$\frac{R_t^+}{R_t^-} = \frac{1 - \mathcal{P}_t}{\mathcal{P}_t}, t \geq 2$$

2.4 Fair pricing condition

The martingale condition derived in the previous section only sets the value of the ratio R_t^+/R_t^- at each auction t. Another condition is required to derive the values of R_t^+ and of R_t^- and therefore, to obtain the expression for the immediate and the permanent impact.

- **The fair pricing condition** states that for any N,

$$\pi_N = \frac{1}{N}\sum_{t=1}^{N} S_t - X_N = 0$$

This implies that we have

$$I_N = \frac{1}{N}\sum_{t=1}^{N} \mathcal{I}_t$$

Assuming that the martingale conditions holds for $t = 1, ..., M$ and the fair pricing condition condition holds for $t = 2, ..., M-1$ leads to a system of $2M-2$ homogeneous equations with $2M-1$ unknowns, so we choose R_1^+ as an undetermined constant.

Proposition 2. *The system of martingale conditions and fair pricing conditions has the solution*

$$R_t^+ = \frac{1}{t}\frac{1-\mathcal{P}_t}{\mathcal{P}_t}\frac{1}{\mathcal{P}_1\mathcal{P}_2...\mathcal{P}_{t-1}}R_1^+, t \geq 2$$
$$R_t^- = \frac{1}{t}\frac{1}{\mathcal{P}_1\mathcal{P}_2...\mathcal{P}_{t-1}}R_1^+, t \geq 2$$

that can be rewritten as :

$$R_t^+ = \frac{1}{t}\frac{p_t}{\sum_{i\geq t+1} p_i}\frac{1-p_1}{\sum_{i\geq t} p_i}R_1^+$$

considering the fact that :

$$\sum_i p_i = 1$$

Corollary 1. *The immediate impact is*

$$\mathcal{I}_t = S_t - X_0 = R_0^+ + R_1^+\left(1 + \sum_{k=2}^{t-1}\frac{1}{k}\frac{1-\mathcal{P}_k}{\mathcal{P}_k}\frac{1}{\mathcal{P}_1...\mathcal{P}_{k-1}}\right), t \geq 2$$

Since $X_N - X_0 = X_N - S_N + S_N - X_0 = \mathcal{I}_N - R_N^-$, we have

Corollary 2. *The permanent impact is*

$$I_N = R_0^+ + R_1^+\left(1 + \sum_{k=2}^{N-1}\frac{1}{k}\frac{1-\mathcal{P}_k}{\mathcal{P}_k}\frac{1}{\mathcal{P}_1...\mathcal{P}_{k-1}} - \frac{1}{N}\frac{1}{\mathcal{P}_1...\mathcal{P}_{N-1}}\right), N \geq 2$$

2.5 Dependence on the metaorder size distribution

According to [Farmer et al., 2013] which cite a host of other relevant studies, there is a considerable evidence that - in the large size limit and for most major equity markets - the metaorder size V is distributed as $\mathbb{P}(V > v) \sim v^{-\beta}$, with $\beta \approx 1.5$.

- $\mathbb{P}(V > v-1) - \mathbb{P}(V > v) \sim v^{-\beta}\left(\left(1 - \frac{1}{v}\right)^{-\beta} - 1\right) \approx \frac{\beta}{v^{\beta+1}}$: an exact zeta distribution for all $n \geq 1$ is considered,

$$p_n = \frac{1}{\zeta(\beta+1)}\frac{1}{n^{\beta+1}}$$

so that

$$\mathcal{P}_t = \frac{\zeta(1+\beta, t+1)}{\zeta(1+\beta, t)} \approx \frac{\displaystyle\int_{t+1}^{+\infty} \frac{\mathrm{d}x}{x^{\beta+1}}}{\displaystyle\int_{t}^{+\infty} \frac{\mathrm{d}x}{x^{\beta+1}}} = \left(1+\frac{1}{t}\right)^{-\beta} \sim 1 - \frac{\beta}{t};$$

- As a consequence,

$$R_t^+ = \frac{1}{t^{2+\beta}} \frac{\zeta(1+\beta)}{\zeta(1+\beta, t)\zeta(1+\beta, t+1)} R_1^+ \sim \frac{1}{t^{2-\beta}};$$

- Thus, the immediate impact $\mathcal{I}_t$ behaves asymptotically for large t as

$$\mathcal{I}_t \sim \begin{cases} t^{\beta-1} & for\, \beta \neq 1 \\ log(t+1) & for\, \beta = 1 \end{cases};$$

- Recalling the fact that the fair pricing gives $I_N = \frac{1}{N}\sum_{t=1}^{N} \mathcal{I}_t$, there holds

$$I_N \sim \frac{1}{N}\int_0^N x^{\beta-1}\,\mathrm{d}x = \frac{1}{\beta} N^{\beta-1}.$$

Finally

Proposition 3.

$$\frac{I_N}{\mathcal{I}_N} = \frac{1}{\beta}$$

for further use, one can observe that for a value of $\beta = 1.5$, we have $\frac{I_N}{\mathcal{I}_N} = \frac{2}{3}$.

3 Definitions, Algorithm and Market Impact measures

3.1 Basic Definitions

Some basic concepts, and the algorithmic definition of a metaorder, are introduced here.

Definition 1. *A **(passive) limit order** is an order that sets the maximum or minimum price at which an agent is willing to buy or sell a given quantity of a particular stock.*

Definition 2. *An **aggressive limit order** is one that instantaneously removes liquidity from the order book by triggering a transaction. An aggressive order crosses the Bid–Ask spread. In other words an aggressive buy order will be placed on the ask, and an aggressive sell order will be placed on the bid.*

Loosely speaking a **metaorder** is a large trading order that is split into small pieces and executed incrementally. In order to perform rigorous statistical analyses, a more specific and precise definition of a metaorder is required, and given in Definition 3 below:

Definition 3. *A* ***metaorder*** *is a series of orders sequentially executed during the same day and having those same attributes:*

- *agent i.e. a participant on the market (an algorithm, a trader...);*
- *product id i.e. a financial instrument (a share, an option...);*
- *direction (buy or sell);*

The advantage of adopting such a definition is that it is no longer necessary to work directly on raw metaorder data. Indeed, a series of orders executed by the same actor on the same product on the financial market will behave like a metaorder and therefore can be considered as such.

AGENT	PRODUCT_ID	DIRECTION	TIMESTAMP	CURRENCY	EXEC_PRICE	EXEC_QTY	PRICE_TYPE	ORDER_TYPE
agent01	product01	Sell	01-APR-16 16:52:00.536612	EUR	87,5	127	Aggressive	Limit
agent05	product11	Buy	01-APR-16 12:03:48.203844	EUR	38,475	393	Aggressive	Limit
agent01	product01	Sell	01-APR-16 13:33:09.566350	EUR	87,4	64	Aggressive	Limit
agent01	product01	Sell	01-APR-16 16:52:00.536641	EUR	87,5	52	Aggressive	Limit
agent04	product20	Sell	01-APR-16 09:46:08.135532	EUR	4,128	570	Aggressive	Limit
agent11	product13	Sell	01-APR-16 15:10:32.279258	EUR	15,2	459	Aggressive	Limit
agent11	product02	Sell	01-APR-16 13:39:30.559504	EUR	4,263	2000	Aggressive	Limit
agent05	product07	Sell	01-APR-16 16:00:29.447697	EUR	58,5	173	Aggressive	Limit
agent05	product11	Buy	01-APR-16 16:34:49.039137	EUR	39,08	374	Aggressive	Limit
agent11	product09	Buy	01-APR-16 13:03:01.846048	EUR	17,005	500	Aggressive	Limit
agent11	product04	Sell	01-APR-16 15:14:18.662973	EUR	19,23	746	Aggressive	Limit
agent11	product02	Buy	01-APR-16 16:12:15.194178	EUR	4,194	942	Aggressive	Limit
agent05	product07	Sell	01-APR-16 14:23:20.168576	EUR	58,32	8	Aggressive	Limit
agent11	product04	Buy	01-APR-16 09:44:28.440733	EUR	19,535	500	Aggressive	Limit
agent01	product01	Sell	01-APR-16 13:33:09.566321	EUR	87,4	100	Aggressive	Limit
agent05	product11	Buy	01-APR-16 12:07:52.669195	EUR	38,45	404	Aggressive	Limit

Figure 2: *Sample of the history of orders* before *the identification process of the algorithm. Each color corresponds to a different metaorder according to the methodology introduced in 3.*

AGENT	PRODUCT_ID	DIRECTION	TIMESTAMP	CURRENCY	EXEC_PRICE	EXEC_QTY	PRICE_TYPE	ORDER_TYPE
agent01	product01	Sell	01-APR-16 13:33:09.566321	EUR	87,4	100	Aggressive	Limit
agent01	product01	Sell	01-APR-16 13:33:09.566350	EUR	87,4	64	Aggressive	Limit
agent01	product01	Sell	01-APR-16 16:52:00.536612	EUR	87,5	127	Aggressive	Limit
agent01	product01	Sell	01-APR-16 16:52:00.536641	EUR	87,5	52	Aggressive	Limit
agent04	product20	Sell	01-APR-16 09:46:08.135532	EUR	4,128	570	Aggressive	Limit
agent05	product07	Sell	01-APR-16 14:23:20.168576	EUR	58,32	8	Aggressive	Limit
agent05	product07	Sell	01-APR-16 16:00:29.447697	EUR	58,5	173	Aggressive	Limit
agent05	product11	Buy	01-APR-16 12:03:48.203844	EUR	38,475	393	Aggressive	Limit
agent05	product11	Buy	01-APR-16 12:07:52.669195	EUR	38,45	404	Aggressive	Limit
agent05	product11	Buy	01-APR-16 16:34:49.039137	EUR	39,08	374	Aggressive	Limit
agent11	product09	Buy	01-APR-16 13:03:01.846048	EUR	17,005	500	Aggressive	Limit
agent11	product04	Buy	01-APR-16 09:44:28.440733	EUR	19,535	500	Aggressive	Limit
agent11	product04	Sell	01-APR-16 15:14:18.662973	EUR	19,23	746	Aggressive	Limit
agent11	product13	Sell	01-APR-16 15:10:32.279258	EUR	15,2	459	Aggressive	Limit
agent11	product02	Sell	01-APR-16 13:39:30.559504	EUR	4,263	2000	Aggressive	Limit
agent11	product02	Buy	01-APR-16 16:12:15.194178	EUR	4,194	942	Aggressive	Limit

Figure 3: *Sample of the history of orders* after *the identification process of the algorithm. Each color corresponds to a different metaorder according to the methodology introduced in 3.*

Let us mention that our study is not limited only to metaorders executed on a same market, due to the fact that an instrument can be simultaneously traded on several markets. Besides, orders executed during the same second are aggregated in order to avoid time stamping issues: the quantities are summed up, the local VWAP is set as the execution price, and the TIMESTAMP of the last order is retained for all orders during the same second. One can see below how those simplifications reduce the complexity of Figure 3:

AGENT	PRODUCT_ID	DIRECTION	TIMESTAMP	CURRENCY	EXEC_PRICE	EXEC_QTY	PRICE_TYPE	ORDER_TYPE
agent01	product01	Sell	01-APR-16 13:33:09.566350	EUR	87,4	164	Aggressive	Limit
agent01	product01	Sell	01-APR-16 16:52:00.536641	EUR	87,5	179	Aggressive	Limit
agent04	product20	Sell	01-APR-16 09:46:08.135532	EUR	4,128	570	Aggressive	Limit
agent05	product07	Sell	01-APR-16 14:23:20.168576	EUR	58,32	8	Aggressive	Limit
agent05	product07	Sell	01-APR-16 16:00:29.447697	EUR	58,5	173	Aggressive	Limit
agent05	product11	Buy	01-APR-16 12:03:48.203844	EUR	38,475	393	Aggressive	Limit
agent05	product11	Buy	01-APR-16 12:07:52.669195	EUR	38,45	404	Aggressive	Limit
agent05	product11	Buy	01-APR-16 16:34:49.039137	EUR	39,08	374	Aggressive	Limit
agent11	product09	Buy	01-APR-16 13:03:01.846048	EUR	17,005	500	Aggressive	Limit
agent11	product04	Buy	01-APR-16 09:44:28.440733	EUR	19,535	500	Aggressive	Limit
agent11	product04	Sell	01-APR-16 15:14:18.662973	EUR	19,23	746	Aggressive	Limit
agent11	product13	Sell	01-APR-16 15:10:32.279258	EUR	15,2	459	Aggressive	Limit
agent11	product02	Sell	01-APR-16 13:39:30.559504	EUR	4,263	2000	Aggressive	Limit
agent11	product02	Buy	01-APR-16 16:12:15.194178	EUR	4,194	942	Aggressive	Limit

Figure 4: *Sample of the history of the orders executed corresponding to the previous figure after merging the orders. The yellow metaorder which was previously a metaorder of length $N = 4$ is now a metaorder of length $N = 2$.*

3.2 Algorithmic procedure

To carry out the study of market impact, the algorithm is trained and run on a proprietary database. Figure 2 in the previous section represents the database in its initial state, as the input of the algorithm. Figure 3 shows an intermediate state of the data extracted from the initial data, during the metaorder reconstruction phase. Figure 4 displays the data in their final state, directly exploitable for statistical studies. Note that the recovery and cleaning of market data are done simultaneously during the metaorder reconstruction.

3.3 Market Impact definitions

We adopt the same framework as that presented in [Bacry et al., 2015]. Let Ω be the set of metaorders under scrutiny, that is, metaorders that are fully executed during a single market session, and pick $\omega \in \Omega$ executed on an instrument S and during a given day d. Its execution starts at some time $t_0(\omega)$ and ends the same day at time $t_0(\omega) + T(\omega)$. Thus $T(\omega)$ represents the duration of the metaorder. Denote by $Q(\omega)$ and $\mathrm{N}(\omega)$ respectively the number of shares and the number of orders that have been executed during the life cycle of the metaorder ω. Hence $Q(\omega)$ is the size and $N(\omega)$ the length of ω. Let $V(\omega)$ be the volume traded the same day d on the instrument S (summed over all European trading venues) between time t_0 and $t_0 + T$, i.e. during the life cycle of ω. The sign of ω will be noted $\epsilon(\omega)$ (i.e. $\epsilon = 1$ for a buy order and $\epsilon = -1$ for a sell order). Clearly, all the quantities introduced in this section depend on ω. For the sake of simplicity, we chose to omit this dependence whenever there is no ambiguity and will often write T, N, Q, V, ϵ instead of $T(\omega)$, $N(\omega)$, $Q(\omega)$, $V(\omega)$, $\epsilon(\omega)$.

The **market impact curve** of a metaorder ω quantifies the magnitude of the relative price variation between the starting time of the metaorder t_0 and the current time $t > t_0$. Let $\mathcal{I}_t(\omega)$ be a proxy for the realized price variation between time t_0 and time $t_0 + t$. In line with many authors ([Almgren et al., 2005], [Bershova and Rakhlin, 2013], [Bacry et al., 2015]), we use the *return proxy* defined by

$$\mathcal{I}_t = \frac{P_t - P_{t_0}}{P_{t_0}}, \tag{1}$$

where P_t represents the execution price of the financial instrument S during the *execution part* of the metaorder, and the mid-price of S during its *relaxation part*, i.e. the decay of the impact, which starts when the metaorder has been fully executed. This estimation relies on the assumption that the *exogenous market moves* W_t will cancel out once averaged, i.e. as a random variable, W_t should have finite variance and basically satisfy $\mathbb{E}(\epsilon(\omega)W_t(\omega)) = 0$.

One can thus write

$$\epsilon(\omega)\mathcal{I}_t(\omega) = \eta_t(\omega) + \epsilon(\omega)W_t(\omega), \tag{2}$$

where $\eta_t(\omega)$ represents the market impact curve and $W_t(\omega)$, the exogenous variation of the price corresponding to the relative price move that would have occurred if the metaorder had not been sent to the market.

4 Empirical study

4.1 Notations

Notation	Definition
ω	a metaorder
$S(\omega)$	financial instrument of the metaorder ω
$d(\omega)$	execution day of the metaorder ω
$t_0(\omega)$	start time of the metaorder ω
$T(\omega)$	duration of the metaorder ω
$N(\omega)$	length of the metaorder ω
$Q(\omega)$	size of the metaorder ω
$V(\omega)$	volume traded on the day $d(\omega)$ on the instrument $S(\omega)$ during $[t_0(\omega), t_0(\omega) + T(\omega)]$
$\epsilon(\omega)$	sign of the metaorder ω
$P(\omega)$	price of $S(\omega)$
$F(\omega)$	future price of the domestic index which contains $S(\omega)$
$\beta_{capm}(\omega)$	beta coefficient of $S(\omega)$ estimated from 6 months *close-to-close* data
$\sigma^*(\omega)$	volatility of $S(\omega)$ estimated from 6 months *close-to-close* data
$\sigma(\omega)$	volatility normalized to the duration of the metaorder ω, i.e. $\sigma(\omega) = \dfrac{\sigma^*(\omega) \times \sqrt{T(\omega)}}{\sqrt{252 \times 8.5 \times 3600}}$
Ω	set of all the metaorders identified by the algorithm
$\Omega_n* \subset \Omega$	subset of the metaorders with $N \geq n*$
$\Omega_{n^*,t^*} \subset \Omega$	subset of the metaorders with $N \geq n^*$ and $T \geq t^*$
$\Omega_{n^*,t^*,*} \subset \Omega$	subset of the metaorders with $N \geq n^*$, $T \geq t^*$ and $t_0 + 2T \leq$ closing time
$Q(S, d)$	quantity total traded by all the bank agents on the instrument S on the day d
T^*	reference constant of a duration in seconds
$(Q/V)^*$	reference constant of a participation rate
I^*	reference constant of a positive impact

Table 2: *Notations and definitions*

Remark 1. *The notations T^*, $(Q/V)^*$ and I^* are some constants.*

4.2 Aggressive Limit Orders

The subject of interest of this section is the metaorders generated by aggressive limit orders, that is, limit orders that actually cross the spread in order to trigger an immediate transaction.

4.2.1 Data

- Study period : **1st Jan 2016 - 30th Aug 2016**
- Markets : **Euronext + Xetra European Equity Markets**
- Order types : **Aggressive Limit Orders**
- Filters : **metaorders** $\omega \in \Omega_{10,600,*}$ **i.e. with a length N $\geq$ 10, a duration T $\geq$ 600 s and** $\mathbf{t}_0 + 2T \leq$ **closing time**
- Number of metaorders : **33054**

4.2.1.1 Duration distribution

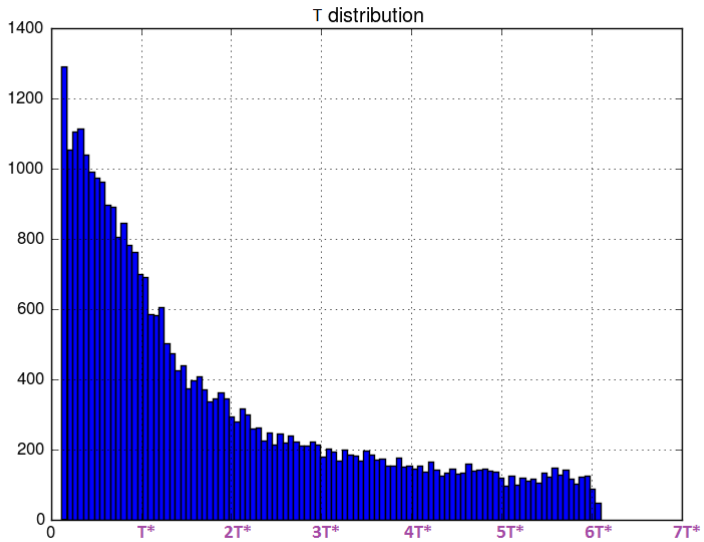


Figure 5: Duration distribution in seconds of the *intraday aggressive metaorders*

mean $1.83 \times T^*$
median $1.24 \times T^*$

One can observe that, in agreement with the intuition, metaorders with shorter length (in time) are proportionally more represented.

4.2.1.2 Length distribution

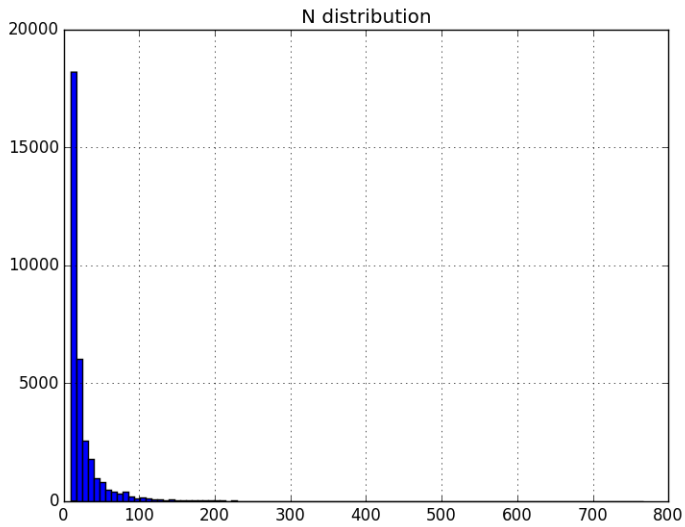


Figure 6: length distribution of the *intraday aggressive metaorders*

mean 25
median 16

As for the duration, shorter metaorders (in length) are also more represented. This observation is not surprising, since the quantities N and T are expected to be highly positively correlated. A log-log scale (Figures 7, 8) gives a more precise idea of the distribution of N: The apparently linear relation suggests that N follows a discrete Pareto distribution, as was already observed in [Bershova and Rakhlin, 2013] and theoretically proven in the market impact model of [Farmer et al., 2013].

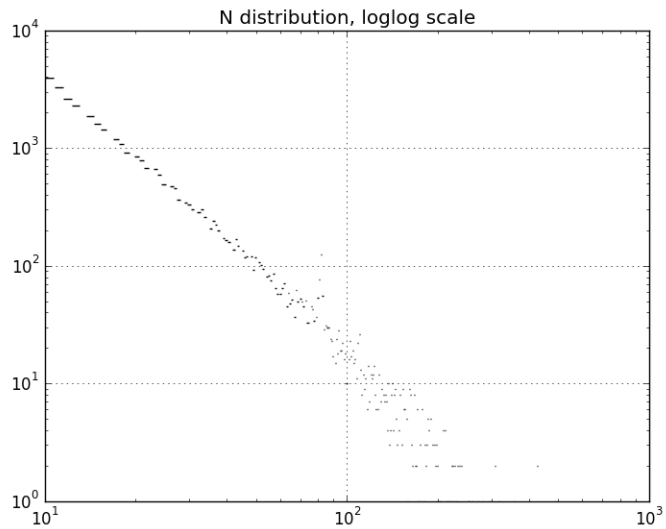
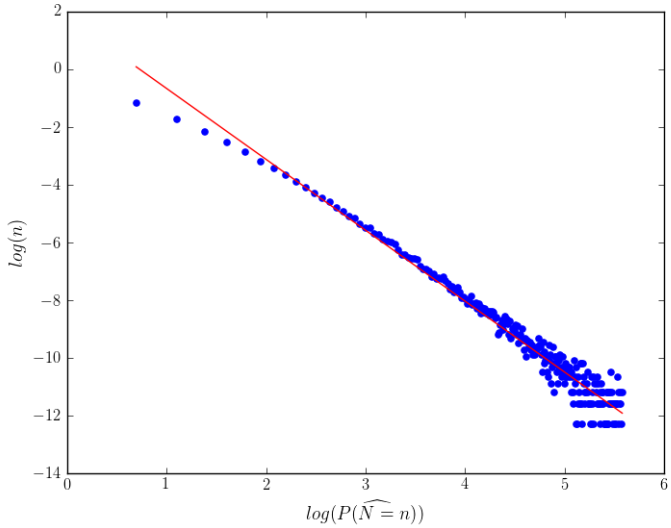


Figure 7: length distribution of the *intraday aggressive metaorders in log-log scale*




```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.967
Model:                            OLS   Adj. R-squared:                  0.967
Method:                 Least Squares   F-statistic:                     7303.
Date:                Tue, 23 May 2017   Prob (F-statistic):          5.59e-186
Time:                        08:19:56   Log-Likelihood:                -136.48
No. Observations:                 250   AIC:                             277.0
Df Residuals:                     248   BIC:                             284.0
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [95.0% Conf. Int.]
------------------------------------------------------------------------------
const          1.7929      0.134     13.395      0.000         1.529     2.057
x1            -2.4566      0.029    -85.460      0.000        -2.513    -2.400
==============================================================================
Omnibus:                       39.718   Durbin-Watson:                   1.584
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               60.738
Skew:                          -0.930   Prob(JB):                     6.47e-14
Kurtosis:                       4.540   Cond. No.                         24.5
==============================================================================
```

Figure 8: Fitting Length distribution of the *intraday aggressive metaorders*.

$\mathbb{P}(\widehat{N=n})$ stands for the natural frequency estimator of the probability $\mathbb{P}(N=n)$. Considering the same definition of β in [Farmer et al., 2013] reminded in section 2.5, we have $\beta = 1.46$ and $\beta_{AllianceBernstein} = 1.56$ in [Bershova and Rakhlin, 2013].

4.2.1.3 Participation rate distribution

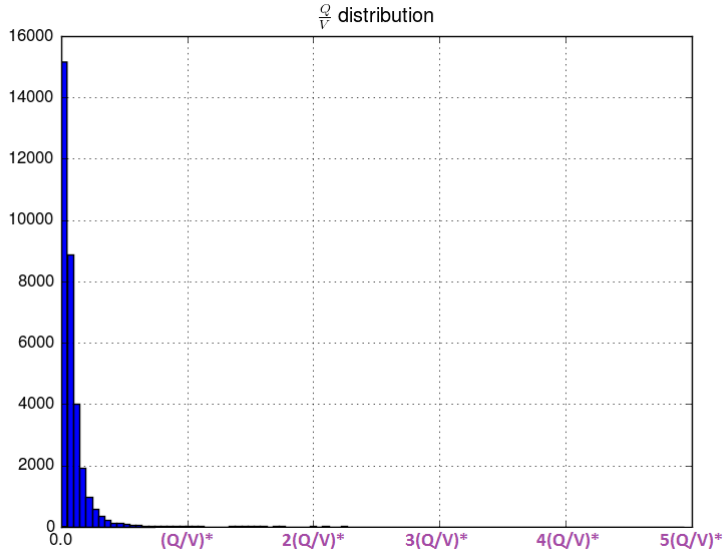


Figure 9: $\frac{Q}{V}$ distribution of the *intraday aggressive metaorders*

mean $0.1 \times (Q/V)^*$
median $0.055 \times (Q/V)^*$

Similar to the case of duration and length, metaorders with smaller participation rates are more represented. Moreover, as one would expect, the quantities N and $\frac{Q}{V}$ are highly positively correlated, a fact that is underlined in Figure 10 below.

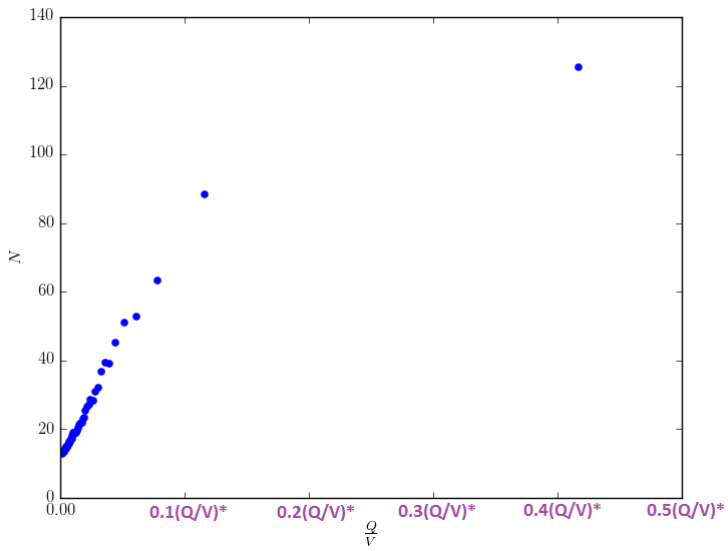


Figure 10: Length N of metaorders as a function of participation rates $\frac{Q}{V}$ in the case of the *aggressive metaorders.*

There is a clear positive correlation between the quantities N and $\frac{Q}{V}$ - with the exception of the last point. However this last point has no real statistical meaning, as it is the result of averaging on the last bucket, thereby mixing very different values of lengths and participation rates on a large range. This singularity appears clearly on the graph.

4.2.2 Market Impact curves

The main results of Section 4.2 are now given, namely, the *market impact curves* for aggressive metaorders. In order to plot explanatory variables, a *bucketing method* is used: Consider for example that one wants to plot y as a function of x, x, y being two arrays of data. One starts by ordering the couple of values (x_i, y_i) by x and then divides the **sorted (by x) distribution** $(x, y)_{sorted}$ into N_{bucket}. Hence one has N_{bucket} subsets of the distribution $(x, y)_{sorted}$, $(x_i, y_i)_{i \in I_1}$, $(x_i, y_i)_{i \in I_2}$, ..., $(x_i, y_i)_{i \in I_{N_{bucket}}}$ and calculates for each bucket I_k the *mean values* $(\overline{x}_k, \overline{y}_k)$. The last step of this average procedure is to plot the points $(\overline{x}_1, \overline{y}_1)$, $(\overline{x}_2, \overline{y}_2)$, ..., $(\overline{x}_{N_{bucket}}, \overline{y}_{N_{bucket}})$.

4.2.2.1 Market Impact Dynamics

To study the dynamics of the market impact, one plots $(\epsilon(\omega)\mathcal{I}_t(\omega))_{\omega \in \Omega, t_0(\omega) \leq t \leq t_0(\omega)+2T(\omega)}$. The first sub-interval $t_0(\omega) \leq t \leq t_0(\omega) + T(\omega)$ corresponds to the execution of the metaorder, whereas the second $t_0(\omega) + T(\omega) \leq t \leq t_0(\omega) + 2T(\omega)$ corresponds to the relaxation.

To perform an extensive statistical analysis over metaorders of varying lengths in physical or volume time, a rescaling in time is necessary, see e.g. [Bacry et al., 2015]. With this convention, all orders are executed during the time interval $[0, 1]$, and the price relaxation occurs in the time interval $[1, 2]$: For each metaorder ω, one considers $[0, 1]$ instead of $[t_0(\omega), t_0(\omega) + T(\omega)]$ $\left([0, 1] = \frac{[t_0(\omega), t_0(\omega) + T(\omega)] - t_0(\omega)}{T(\omega)}\right)$ for the *execution part* of ω and $[1, 2]$ instead of $[t_0(\omega) + T(\omega), t_0(\omega) + 2T(\omega)]$ for the *relaxation part* of ω, and then averages using the *bucketing method* previously described on the time-rescaled volume quantities.

The time variable $t \in [0, 1]$ in Figure 11 is actually the volume time i.e. the ratio between the part of the metaorder already executed at the time of the observation and the total size of the metaorder - of course, at the end of the execution part this quantity is always equal to 1.

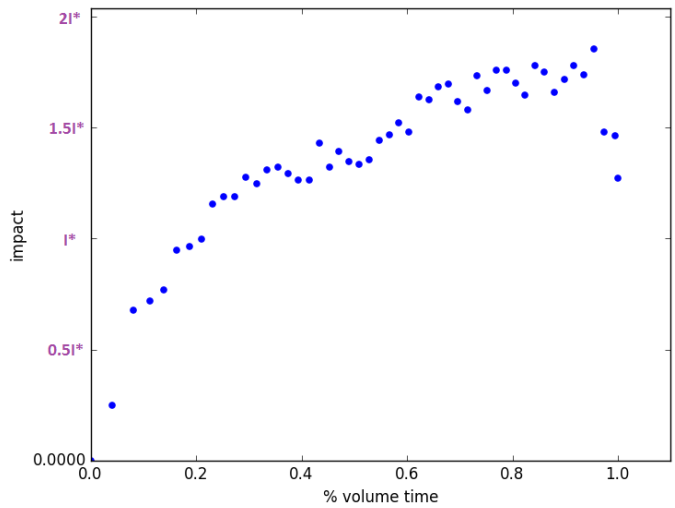


Figure 11: market impact dynamics during the *execution part* in the case of the *aggressive metaorders*.

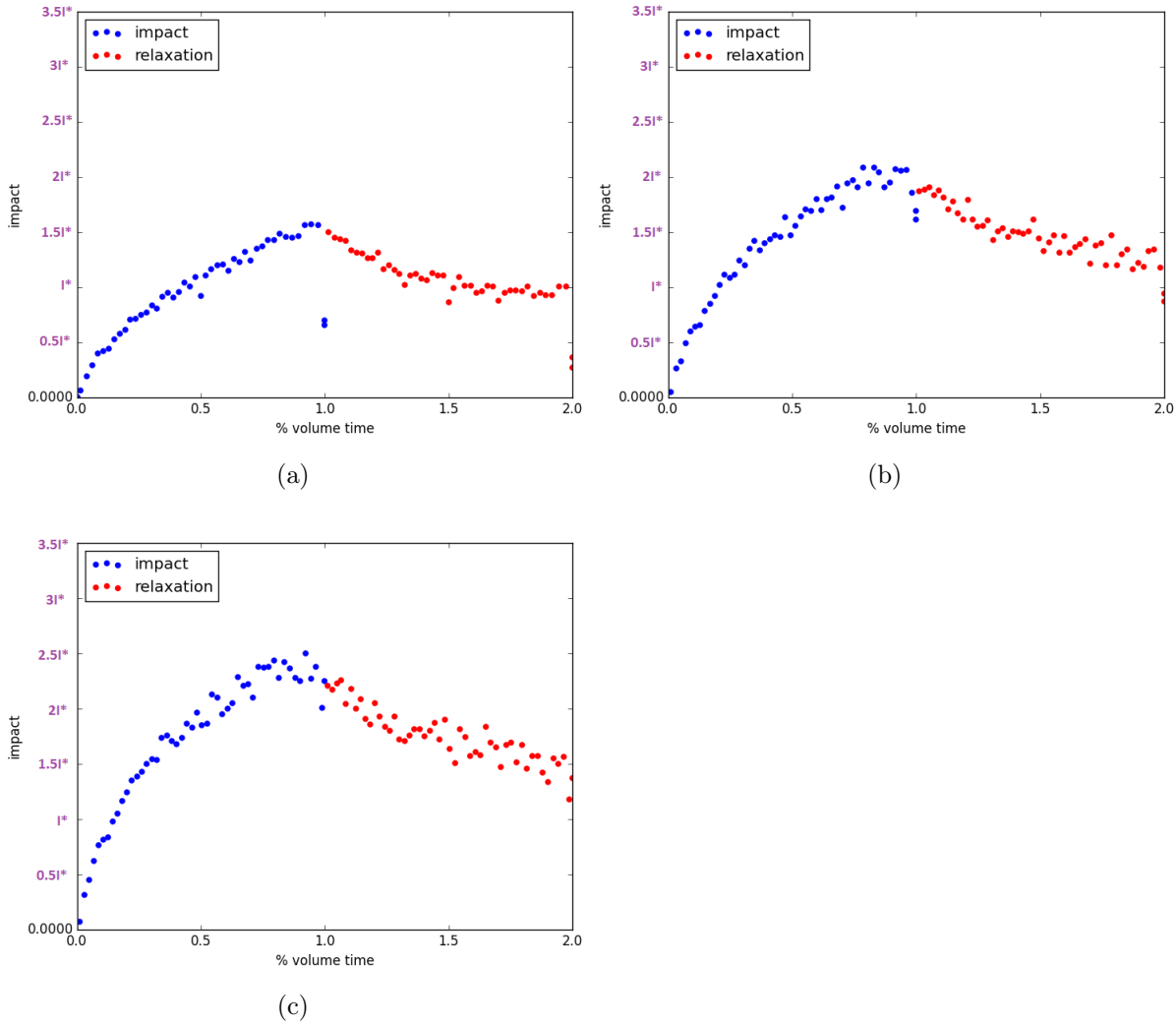


Figure 12: market impact dynamics during the *execution part* and *relaxation part* (a) metaorders of $\Omega_{0,0,*}$, (b) metaorders of $\Omega_{10,0,*}$ (with $N \geq 10$) and (c) metaorders of $\Omega_{20,0,*}$ (with $N \geq 20$).

The analysis yields an increasing, concave market impact curve. The decay observed in the part of the curve points is only a lure and can be explained by the fact that metaorders of smaller lengths are more numerous and associated with lower impacts as shown in Figure 10.

The blue points correspond to execution prices and the red points correspond to mid prices taken at symmetrical times from the end of the metaorder, thus the interval $[1, 2]$ has no physical sense and it is just an obvious choice in order to normalize the shape of the curve. The two isolated points observed in 12a are due to the fact that we have considered all metaorders, so metaorders with a small length, especially those with length $N = 2$ are over-represented in $\%\text{volume} = 1.0$, therefore inducing a bias towards the end of the curve. One can make this artifact vanish when considering metaorders with larger sizes, for which the impact clearly increases (cf. 12b and 12c). Those curves confirm the fact advocated in [Farmer et al., 2013] that the impact is concave and increasing, and that the final impact after the execution is performed relaxes to about two-thirds of the peak impact.

The concave permanent impact and the convex relaxation that follows have already been

underlined by [Bacry et al., 2015]. Also, the relaxation to about two-thirds of the peak impact, predicted by [Farmer et al., 2013], had been observed in [Bershova and Rakhlin, 2013].

On Figures 12a, 12b and 12c, one observes a concave dynamics of the market impact followed by a convex and decreasing relaxation. We note that on the three graphs, as we retain the largest metaorders, we observe higher market impacts: $1.5\times I^*$, $2\times I^*$ and then $2.5\times I^*$ for the temporary market impact. On Figure 12a, one can safely assume that relaxation is complete and stable at a level around I^*. However, on Figure 12c, relaxation seems to only begin its stabilization phase and does not appear quite finished. This rather predictable observation expresses the fact that relaxation takes even longer for larger metaorders.

4.2.2.2 Market Impact as a function of the participation rate

We plot here the market impact observed at the end of the *execution part*, i.e. *the temporary market impact*, as a function of the participation rate $\frac{Q}{V}$.

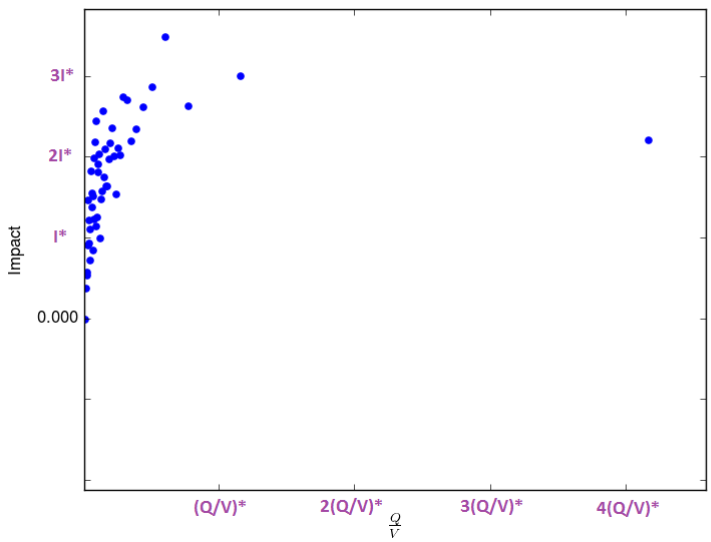
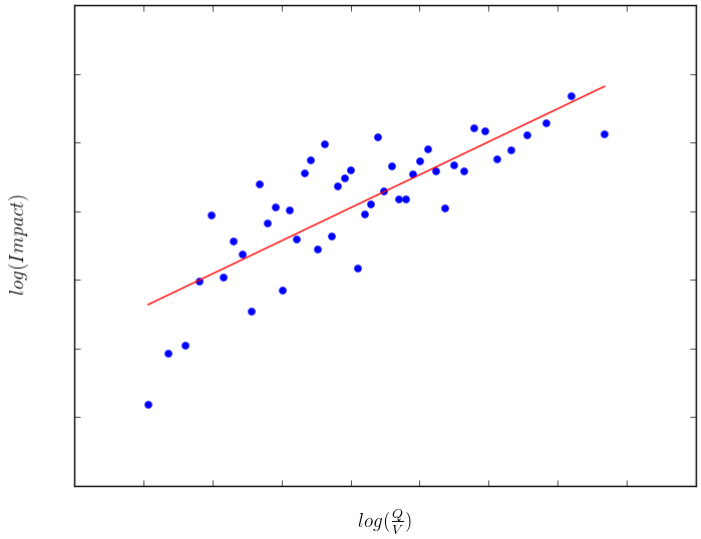


Figure 13: temporary market impact as a function of $\frac{Q}{V}$

This graph also shows an increasing concave curve. The first point corresponds to weak participation rates, which rather suffer from market movements than impose them. Similarly, the last point has no real statistical value because it is the result of averages on the last bucket which contains less data than the previous ones and mixes values of participation rates on a very large range of values.




```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.655
Model:                            OLS   Adj. R-squared:                  0.647
Method:                 Least Squares   F-statistic:                     83.60
Date:                Tue, 23 May 2017   Prob (F-statistic):           9.81e-12
Time:                        15:13:18   Log-Likelihood:                -6.4165
No. Observations:                  46   AIC:                             16.83
Df Residuals:                      44   BIC:                             20.49
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [95.0% Conf. Int.]
------------------------------------------------------------------------------
const         -3.5912      0.315    -11.393      0.000        -4.226    -2.956
x1             0.4808      0.053      9.143      0.000         0.375     0.587
==============================================================================
Omnibus:                        0.350   Durbin-Watson:                   1.739
Prob(Omnibus):                  0.839   Jarque-Bera (JB):                0.270
Skew:                          -0.177   Prob(JB):                        0.874
Kurtosis:                       2.874   Cond. No.                         46.3
==============================================================================
```

Figure 14: temporary market impact as a function of $\frac{Q}{V}$ - power law fitting on buckets

One observes a concave market impact increasing with the participation rate $\frac{Q}{V}$. We fit here with a log-log scale keeping only the positive y-axis plots and neglecting the last point from the Figure 13, a power law to measure the temporary market impact of the *aggressive metaorders*. Following the framework introduced by [Farmer et al., 2013] we obtain two different measurements of $\beta - 1$, one obtained from the distribution of N and the other one from the participation rate $\frac{Q}{V}$: 0.46 and 0.48. In all cases these two values are close to 0.5 and seem to argue in favor of a concave temporary market impact following a square root law also for our *aggressive limit metaorders*.

4.2.3 Conclusion

In this section we have observed that *aggressive metaorders* have behaviors often well described in the scientific literature such as the concave increasing shape of the market impact and the convex decreasing shape of the relaxation. Those issues have been mainly discussed in [Bershova and Rakhlin, 2013], [Farmer et al., 2013], [Bacry et al., 2015], [Gatheral, 2010] and [Gatheral et al., 2011]. Besides we notice that the impact relaxed to about two-thirds of peak impact, and thus the price moves of our *aggressive metaorders* could be well described by a *VWAP AFS model* presented in [Gatheral et al., 2011].

4.3 Passive limit orders

The subject of interest of this section is the metaorders generated by *passive limit orders*, mainly through order execution strategies.

4.3.1 Data

- Study period : **1st Jan 2016 - 30th Aug 2016**
- Markets : **All European Equity Markets**
- Order types : **All Orders ($> 90\%$ passive limit orders)**
- Filters : **metaorders $\omega \in \Omega_{5,600,*}$ with a length N $\geq$ 5, a duration T $\geq$ 600 s and $\mathbf{t}_0 + 2T \leq$ closing time**
- Number of metaorders : **16581**

4.3.1.1 Duration distribution

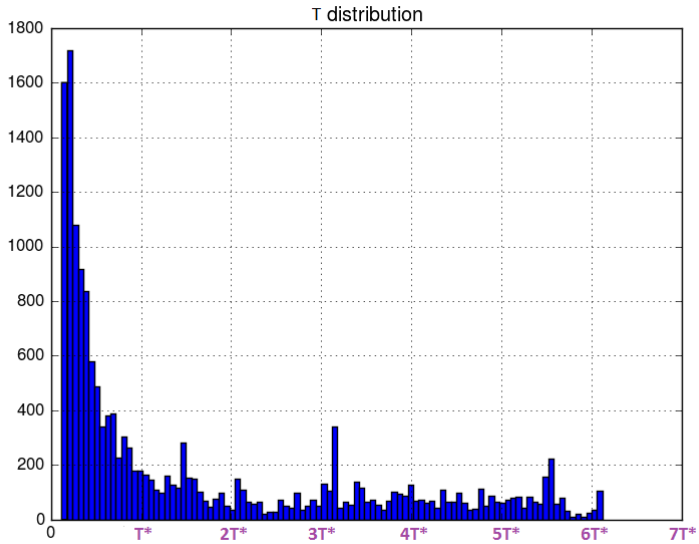


Figure 15: Duration distribution in seconds of the *intraday execution metaorders*

mean $1.63 \times T^*$
median $0.71 \times T^*$

As expected, and previously shown for *aggressive metaorders* (cf. Figure 5), one observes again (Figure 15) that, the shorter the metaorders are, the more they are represented. Contrarily to the case of the aggressive metaorders, the distribution is seen to have a fatter tail, an indication of the 'intelligence' of the execution strategies.

4.3.1.2 Length distribution

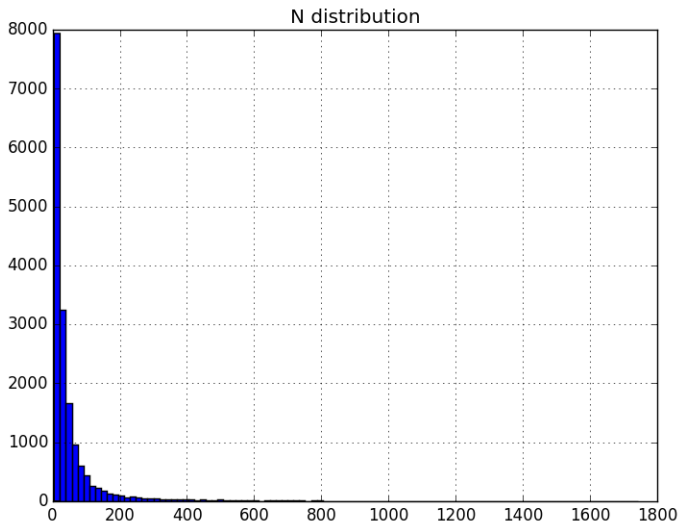


Figure 16: length distribution of the *intraday execution metaorders*

mean 54
median 24

The shape of the distribution in Figure 16 is very similar to the one observed for the *aggressive metaorders* in Figure 6.

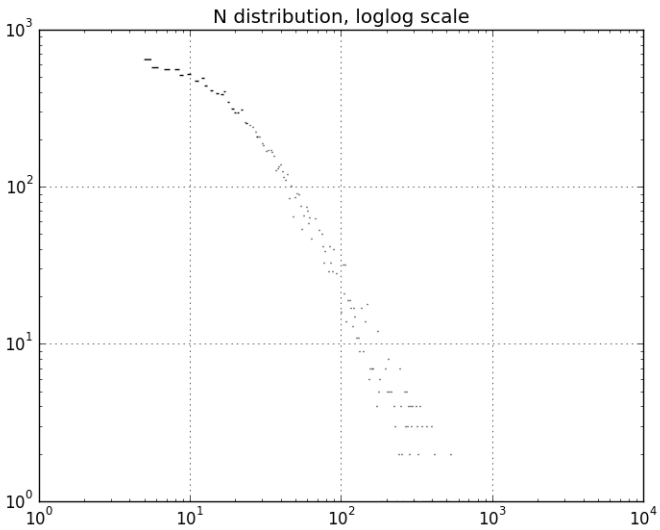


Figure 17: length distribution of the *intraday execution metaorders with a log-log scale*

Again, a linear relation in log-log scale seems clear. This confirms the relevancy of the discrete Pareto distribution independently of the nature (*passive, aggressive...) of the orders.*

4.3.1.3 Participation rate distribution

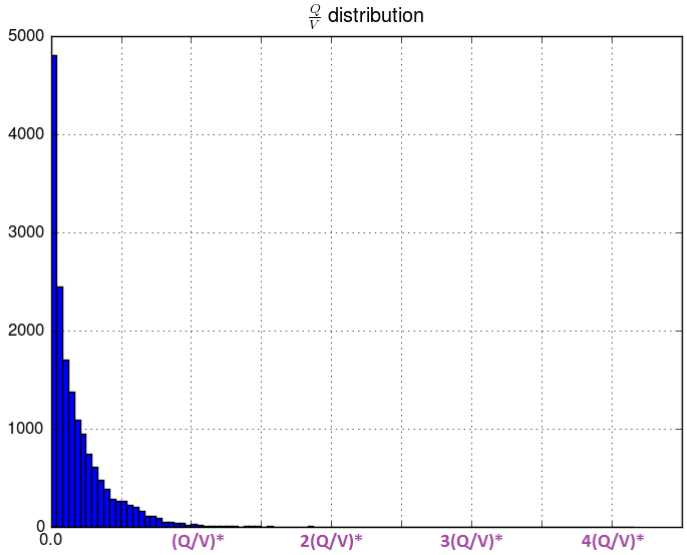


Figure 18: $\dfrac{Q}{V}$ distribution of the *intraday execution metaorders*

$$\textbf{mean } 0.18 \times (Q\backslash V)^*$$
$$\textbf{median } 0.11 \times (Q\backslash V)^*$$

Again the quantities N, T and $\dfrac{Q}{V}$ are highly positively correlated. Besides, it is not surprising to note that the *execution metaorders* have participation rates during their life cycle bigger than those observed for the *aggressive metaorders* (Figure 9).

4.3.2 Market Impact Dynamics

The market impact study proceeds along the same lines as those presented in 4.2.2.1. Hence, for plotting $(\epsilon(\omega)\mathcal{I}_t(\omega))_{\omega\in\Omega, t_0(\omega)\leq t\leq t_0(\omega)+T(\omega)}$, one has to consider its rescaled time version $(\epsilon(\omega)\mathcal{I}_s(\omega))_{\omega\in\Omega, 0\leq s\leq 1}$, i.e. for each metaorder ω we consider $[0,1]$ instead of $[t_0(\omega), t_0(\omega)+T(\omega)]$ such as $[0,1] = \dfrac{[t_0(\omega), t_0(\omega)+T(\omega)] - t_0(\omega)}{T(\omega)}$.

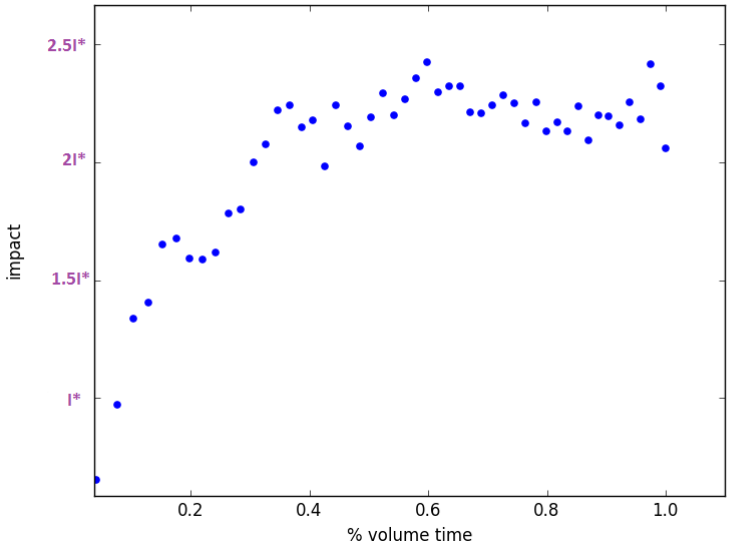


Figure 19: market impact dynamics during the *execution part* in the case of the *intraday execution metaorders*.

One observes a positive increasing and concave market curve. Note that the impacts are more important than those obtained for the *aggressive metaorders* (Figure 11). This is essentially due to the fact that the metaorders involved in an execution strategy are generally bigger than aggressive ones.

4.3.3 Market Impact Models

4.3.3.1 Methodology

In this section, we calibrate a phenomenological model for the market impact of *execution strategies*.

Starting from all values of $\mathcal{I}_t = \dfrac{P_t - P_0}{P_0}$, we try to minimize the noise in order to infer efficiently a market impact model.

$\mathcal{I}_t$ being a *return value*, it seems natural to consider $\mathcal{F}_t = \dfrac{F_t - F_0}{F_0}$ as an explanatory variable due to the fact that the Future moves are often highly positively correlated with the Spot moves. So we are thinking of inferring a market impact model considering the residual terms. And then we shape our model including incrementally relevant explanatory variables such as the volatility σ, the participation rate term $\dfrac{Q}{V}$.

4.3.3.2 Models

- Model M1 : $\mathcal{I}_t = \beta \times \mathcal{F}_t + cst$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                returns   R-squared:                       0.219
Model:                            OLS   Adj. R-squared:                  0.219
Method:                 Least Squares   F-statistic:                     4657.
Date:                Mon, 26 Sep 2016   Prob (F-statistic):               0.00
Time:                        13:39:14   Log-Likelihood:                 57035.
No. Observations:               16581   AIC:                        -1.141e+05
Df Residuals:                   16579   BIC:                        -1.140e+05
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [95.0% Conf. Int.]
------------------------------------------------------------------------------
Intercept     -0.0003   6.03e-05     -5.038      0.000        -0.000    -0.000
returns_fut    0.5945      0.009     68.239      0.000         0.577     0.612
==============================================================================
Omnibus:                     4911.773   Durbin-Watson:                   1.747
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           566051.030
Skew:                          -0.328   Prob(JB):                         0.00
Kurtosis:                      31.616   Cond. No.                         145.
==============================================================================
```

Figure 20: results *execution strategies* Model M1
The Futures moves explain 22% of the Stock moves.

- Model M2 : $\mathcal{I}_t = \beta \times \beta_{capm} \times \mathcal{F}_t + cst$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                returns   R-squared:                       0.252
Model:                            OLS   Adj. R-squared:                  0.252
Method:                 Least Squares   F-statistic:                     5593.
Date:                Mon, 26 Sep 2016   Prob (F-statistic):               0.00
Time:                        13:39:14   Log-Likelihood:                 57392.
No. Observations:               16581   AIC:                        -1.148e+05
Df Residuals:                   16579   BIC:                        -1.148e+05
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [95.0% Conf. Int.]
------------------------------------------------------------------------------
Intercept     -0.0003    5.9e-05     -4.422      0.000        -0.000    -0.000
beta_fut       0.7197      0.010     74.786      0.000         0.701     0.739
==============================================================================
Omnibus:                     5190.039   Durbin-Watson:                   1.740
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           656820.406
Skew:                          -0.416   Prob(JB):                         0.00
Kurtosis:                      33.822   Cond. No.                         163.
==============================================================================
```

Figure 21: results *execution strategies* Model M2
When normalizing the Futures moves using stock's β_{capm}, one explains 3% more of the stock returns.

- Model M3 : $\mathcal{I}_t = \beta \times \beta_{capm} \times \mathcal{F}_t + k \times \epsilon \times \dfrac{Q}{V} + cst$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                returns   R-squared:                       0.265
Model:                            OLS   Adj. R-squared:                  0.265
Method:                 Least Squares   F-statistic:                     2985.
Date:                Mon, 26 Sep 2016   Prob (F-statistic):               0.00
Time:                        13:39:14   Log-Likelihood:                 57532.
No. Observations:               16581   AIC:                        -1.151e+05
Df Residuals:                   16578   BIC:                        -1.150e+05
Df Model:                           2
Covariance Type:            nonrobust
==================================================================================
                     coef    std err          t      P>|t|      [95.0% Conf. Int.]
----------------------------------------------------------------------------------
Intercept         -0.0003   5.87e-05     -5.716      0.000        -0.000    -0.000
beta_fut           0.7106      0.010     74.346      0.000         0.692     0.729
sgn_part_rate      0.0175      0.001     16.807      0.000         0.015     0.020
==============================================================================
Omnibus:                     5214.671   Durbin-Watson:                   1.746
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           699542.373
Skew:                          -0.403   Prob(JB):                         0.00
Kurtosis:                      34.810   Cond. No.                         163.
==============================================================================
```

Figure 22: results *execution strategies* Model M3

Adding the signed participation rate one observes an improvement of 1.3% in the explanation.

- Model M4 : $\mathcal{I}_t = \beta \times \beta_{capm} \times \mathcal{F}_t + k \times \epsilon \times \sqrt{\dfrac{Q}{V}} + cst$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                returns   R-squared:                       0.269
Model:                            OLS   Adj. R-squared:                  0.269
Method:                 Least Squares   F-statistic:                     3054.
Date:                Mon, 26 Sep 2016   Prob (F-statistic):               0.00
Time:                        13:39:14   Log-Likelihood:                 57583.
No. Observations:               16581   AIC:                        -1.152e+05
Df Residuals:                   16578   BIC:                        -1.151e+05
Df Model:                           2
Covariance Type:            nonrobust
==================================================================================
                     coef    std err          t      P>|t|      [95.0% Conf. Int.]
----------------------------------------------------------------------------------
Intercept         -0.0004   5.86e-05     -6.459      0.000        -0.000    -0.000
beta_fut           0.6989      0.010     73.002      0.000         0.680     0.718
sqrt_part_rate     0.0061      0.000     19.625      0.000         0.005     0.007
==============================================================================
Omnibus:                     5262.102   Durbin-Watson:                   1.754
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           714666.686
Skew:                          -0.418   Prob(JB):                         0.00
Kurtosis:                      35.152   Cond. No.                         164.
==============================================================================
```

Figure 23: results *execution strategies* Model M4

Using the square root gives better results, one observes an improvement of 0.4% in the explanation.

- Model M5 : $\mathcal{I}_t = \beta \times \beta_{capm} \times \mathcal{F}_t + k \times \epsilon \times \sigma \times \sqrt{\frac{Q}{V}} + cst$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                returns   R-squared:                       0.271
Model:                            OLS   Adj. R-squared:                  0.271
Method:                 Least Squares   F-statistic:                     3086.
Date:                Mon, 26 Sep 2016   Prob (F-statistic):               0.00
Time:                        13:39:14   Log-Likelihood:                 57606.
No. Observations:               16581   AIC:                        -1.152e+05
Df Residuals:                   16578   BIC:                        -1.152e+05
Df Model:                           2
Covariance Type:            nonrobust
==================================================================================================
                         coef    std err          t      P>|t|      [95.0% Conf. Int.]
--------------------------------------------------------------------------------------------------
Intercept             -0.0004   5.85e-05     -6.297      0.000        -0.000    -0.000
beta_fut               0.6846      0.010     70.945      0.000         0.666     0.703
vol_sqrt_part_rate     0.8716      0.042     20.825      0.000         0.790     0.954
==============================================================================
Omnibus:                     5562.585   Durbin-Watson:                   1.756
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           713625.222
Skew:                          -0.556   Prob(JB):                         0.00
Kurtosis:                      35.120   Cond. No.                         719.
==============================================================================
```

Figure 24: results *execution strategies* Model M5
Adding the volatility factor enhances the results, one observes an improvement of 0.2% in the explanation.

We conclude this section by plotting the $\mathcal{R}^2$ coefficient as a function of γ in the following model

$$\mathcal{I}_t = \beta \times \beta_{capm} \times \mathcal{F}_t + k \times \epsilon \times \sigma \times \left(\frac{Q}{V}\right)^{\gamma} + cst$$

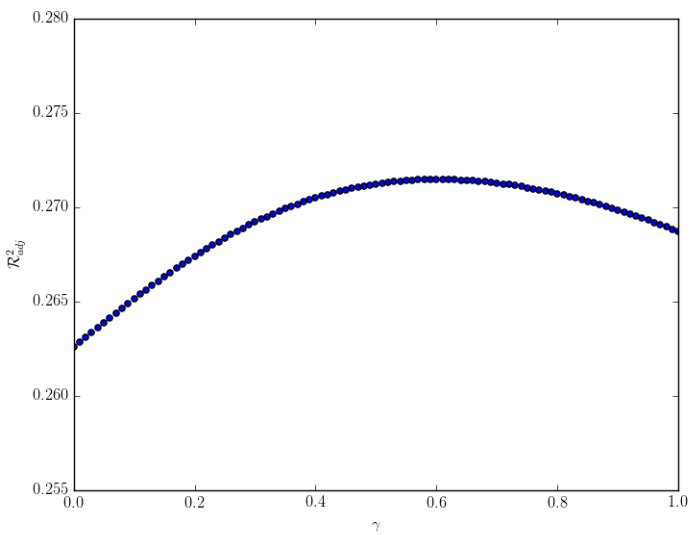


Figure 25: $\mathcal{R}^2$ as a function of γ in the model M5
One notices that the $\mathcal{R}^2$ factor reaches its maximum in $\gamma = 0.60$. We have $\beta = 0.69$, $k = 1.2$ and $\mathcal{R}^2 = 0.271$.

As a remark, let us add that we also tried some multi-factor models but those models do not really perform better than model M5, so we decided to keep a single factor model.

4.3.3.3 Stability

The subject of interest of this section is the stability over time of the following model

$$\mathcal{I}_t = \beta \times \beta_{capm} \times \mathcal{F}_t + k \times \epsilon \times \sigma \times \left(\frac{Q}{V}\right)^{\gamma} + cst$$

In order to do this, we proceed as in the previous section over a 3 month sliding window. The results are summarized in the table below.

period	Nb metaorders	γ	R^2_{adj}	β	k
1 Jan 2016 - 30 Mar 2016	7859	0.42	0.271	0.6295	0.8352
1 Feb 2016 - 30 Apr 2016	7991	0.92	0.240	0.6676	2.4797
1 Mar 2016 - 30 May 2016	7654	0.52	0.221	0.5854	0.8734
1 Apr 2016 - 30 Jun 2016	6786	0.87	0.270	0.7924	2.9644
1 May 2016 - 30 Jul 2016	5505	0.42	0.336	0.7901	0.8829
1 Jun 2016 - 30 Aug 2016	3709	0.39	0.377	0.8423	0.6801

Table 3: Stability for each period with the optimal power law formula

period	Nb metaorders	γ	R^2_{adj}	β	k
1 Jan 2016 - 30 Mar 2016	7859	0.50	0.271	0.6318	1.1044
1 Feb 2016 - 30 Apr 2016	7991	0.50	0.239	0.6598	0.6226
1 Mar 2016 - 30 May 2016	7654	0.50	0.221	0.5849	0.8149
1 Apr 2016 - 30 Jun 2016	6786	0.50	0.267	0.7801	0.8955
1 May 2016 - 30 Jul 2016	5505	0.50	0.335	0.7970	1.1436
1 Jun 2016 - 30 Aug 2016	3709	0.50	0.376	0.8543	0.9646

Table 4: Stability for each period with a square root formula

One can conclude from these two tables and the graphs just below (Figure 26) that choosing γ close to 0.5 is a reasonable assumption. Also, the values of β and k seem to be more stable with a square root formula. This tends to confirm the commonly agreed upon square root behaviour for the impact of execution strategies.

As a conclusion to this short modelling section, one can observe that significant improvements occur as we shape our model with appropriate factors. As expected, the returns of the Futures price is the most relevant factor.

Moreover, concerning the *execution metaorders*, the participation rate $\frac{Q}{V}$ and the volatility σ seem to be the most relevant factors, and market impact is seen to behave like $\sigma \times \left(\frac{Q}{V}\right)^{\gamma}$ with $\gamma \approx 0.5$.

5 Conclusion

The work presented here is an empirical study of a large set of metaorders executed in the European equity markets. A new algorithmic definition of a metaorder has been proposed. The statistical results show a good agreement with some predictions of the market impact model in [Farmer et al., 2013] in the case of limit orders.

Our study contains two distinct subgroups of orders: a set of aggressive limit orders, and a database of execution strategies predominantly composed of passive limit orders. In both cases,

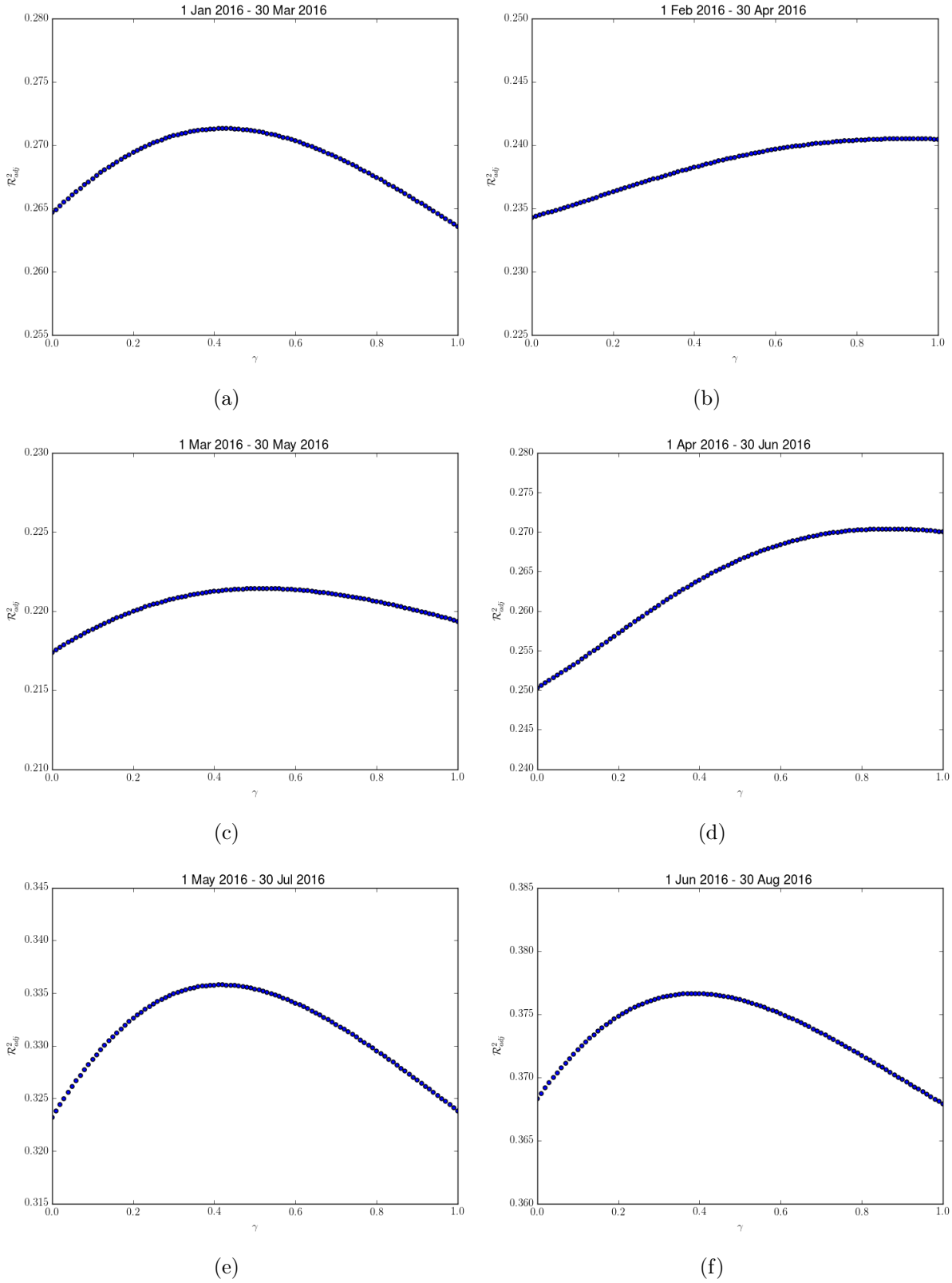


Figure 26: $\mathcal{R}^2$ coefficient as a function of γ for each period in the model M5

the analysis shows that the temporary market impact is increasing and concave, and it also confirms that the length distribution of metaorders follows a Pareto distribution of parameter $\beta + 1$ with $\beta \approx 1.5$ in both cases.

Moreover, in the case of aggressive metaorders, we have been interested in the relaxation phase. A convex, decreasing functional form is obtained, as mentioned in [Bacry et al., 2015] and other related studies. More precisely, the price reversion after the completion of a trade yields a permanent impact is also a square root function of trade volume duration and that its ratio to the maximum impact observed at the last fill is roughly $2/3$ as predicted in the article of [Farmer et al., 2013] and highlighted empirically [Bershova and Rakhlin, 2013].

Finally, in the case of passive limit orders, we have shown that the market impact of metaorders is well described by a square root formula $\sigma\sqrt{\frac{Q}{V}}$, σ being the volatility and $\frac{Q}{V}$, the participation rate.

References

[Almgren et al., 2005] Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct estimation of equity market impact. *Risk*, 18(7):58–62.

[Bacry et al., 2015] Bacry, E., Iuga, A., Lasnier, M., and Lehalle, C.-A. (2015). Market impacts and the life cycle of investors orders. *Market Microstructure and Liquidity*, 1(02):1550009.

[Bershova and Rakhlin, 2013] Bershova, N. and Rakhlin, D. (2013). The non-linear market impact of large trades: Evidence from buy-side order flow. *Quantitative Finance*, 13(11):1759–1778.

[Farmer et al., 2013] Farmer, J. D., Gerig, A., Lillo, F., and Waelbroeck, H. (2013). How efficiency shapes market impact. *Quantitative Finance*, 13(11):1743–1758.

[Gatheral, 2010] Gatheral, J. (2010). No-dynamic-arbitrage and market impact. *Quantitative finance*, 10(7):749–759.

[Gatheral et al., 2011] Gatheral, J., Schied, A., and Slynko, A. (2011). Exponential resilience and decay of market impact. *Econophysics of order-driven markets*, pages 225–236.

[Gomes and Waelbroeck, 2015] Gomes, C. and Waelbroeck, H. (2015). Is market impact a measure of the information value of trades? market response to liquidity vs. informed metaorders. *Quantitative Finance*, 15(5):773–793.

[Moro et al., 2009] Moro, E., Vicente, J., Moyano, L. G., Gerig, A., Farmer, J. D., Vaglica, G., Lillo, F., and Mantegna, R. N. (2009). Market impact and trading profile of hidden orders in stock markets. *Physical Review E*, 80(6):066102.

[Toth et al., 2011] Toth, B., Lemperiere, Y., Deremble, C., De Lataillade, J., Kockelkoren, J., and Bouchaud, J.-P. (2011). Anomalous price impact and the critical nature of liquidity in financial markets. *Physical Review X*, 1(2):021006.