SCUOLA
NORMALE
SUPERIORE

Market impact for large institutional investors: empirical evidences and theoretical models.

Tesi di Perfezionamento
in
Matematica per la Finanza

presentata da
Frédéric Bucci

Supervisor

Prof. Fabrizio Lillo

Contents

Abstract

Advances in technology have deeply changed the way how securities are traded. The introduction of new technologies has enabled exchanges to automate the majority of their trading operations, leading on one side to a considerable cost reductions and on the other side offering a full set of new possibilities for market participants. In parallel to this automation, also brokers, hedge funds, proprietary trading firms, and other market participants have profitted from these new technology approaches for automating a variety of tasks, from optimization of order execution to whole trading strategies. In particular, with this progressive market automation a large amount of data becomes available, representing a unique laboratory where to discover new stylized facts and where to test new financial theories proposed in literature. This possibility has opened new challenges finalized to exploit this information for quantitative research and trading purposes, with a particular focus in market microstructure.

For example, it is commonly accepted that market price moves during the execution of a trade - in average it increases for a buy order and vice versa it decreases for a sell order. This phenomenon, coined as market impact, is clearly a question of great relevance when studying the price formation process and it has also become a major practical issue for brokers, market makers and institutional investors in the design of their optimal trading strategies. Indeed, in order to know whether a trade will be profitable, it is essential to monitor transactions costs, which are directly related to market impact. Measuring and modelling market impact has therefore become a central question of interest both for finance researchers and practitioners with ongoing effort to generate trading model ideas and cost modelling improvements that help with portfolio construction techniques.

In this research strand, one of the most surprising stylized facts is that the market impact of a so called metaorder - a long sequence of orders executed sequentially in the same direction and originated by the same trading decision - is approximately described by a square-root law of the order size, and not linearly as one may have naively expected. In general, public transaction data are not sufficient to perform market impact analysis since from the available information it is not possible to clearly identify the metaorders. This is why much of the work in both the academic and the industrial communities has been done using proprietary datasets. However, from these datasets focusing singularly on a single financial institution at the time it follows that i) the results are specific of the strategy and execution style of the institution, ii) it is not possible to have insights on how the metaorders executed from several institutions interact. For these reasons,

one of the principal aims of this thesis is to investigate the interaction effects on market impact using a data-driven approach based on a rich dataset of metaorders originated by an heterogeneous set of investors in the U.S. equity market. The thesis is organized as follows.

Chapter 1 provides an overview of the most relevant contributions in this thesis.

Chapter 2 introduces the themes and research questions that we address for a general audience. A reader already familiar with the concepts related to the market impact and more in general with the market microstructure field could skip this part.

Chapter 3 presents the ANcerno dataset used through this thesis for the several empirical analysis. It is represented by a rich dataset of heterogeneous institutional investors metaorders traded in the U.S. equity market. We describe some summary statistics of these metaorders introducing the parameters used to characterize their execution.

Chapter 4 describes the first empirical study meant to validate a recent model of market impact based on a dynamical theory of liquidity. We find that the theoretical predictions, based on reaction-diffusion equations in a multiple-time scales framework, are remarkably well borne out by data: a transition from a linear to a square root market impact is observed, as predicted by the theory.

Chapter 5 is devoted to the study of how the square-root law emerges from the interaction between different agents executing metaorders on the same asset. The crowding effects on market impact are investigated and special care is devoted to construct statistical models, which calibrated on data allow us to reproduce very well the different regimes of the empirical market impact curves.

Chapter 6 is focused to shed light on what happens to the price dynamics after the metaorder execution. This is coined as price relaxation and we use several approaches in such a way to clarify the role of the order flow correlation on how the price relaxes after the end of the metaorder both at the intraday and at the multi-day levels. We find that relaxation takes place as soon as the metaorder ends and it continues in the following days with no apparent saturation at any predictable plateau.

Chapter 7 concerns the trading cost associated with the execution of a metaorder which allows to define a natural dimensionless invariant in agreement with the trading invariance principle recently postulated in financial literature. From our empirical evidences it emerges that the trading invariance can be justified from the validity of the square-root law for market impact and from the proportionality between spread and volatility.

Appendix A is devoted to underline that market impact should not be miscontrued as volatility. In particular, the square-root market impact has nothing to do with price diffusion, i.e. that typical price changes grow as the square-root of time. We therefore rationalise empirical findings on market impact and volatility by introducing a simple scaling argument in agreement with data.

Chapters 4 through 7 and Appendix A contain the original contributions of this thesis. Each of them is self-contained and in principle can be read separately.

List of publications

The following articles are part of this thesis:

- Bucci, F., Benzaquen, M., Lillo, F., Bouchaud, J.-P. (2019) *Crossover from linear to square root market impact.* Physical Review Letters, 122(10), 108302.
- Bucci, F., Mastromatteo, I., Eisler, Z., Lillo, F., Bouchaud, J.-P., Lehalle, C.-A. (2019), *Co-impact: crowding effects in institutional trading activity.* Quantitative Finance, 20:2, 193-205, DOI: 10.1080/14697688.2019.1660398.
- Bucci, F., Benzaquen, M., Lillo, F., Lehalle C.-A., Bouchaud, J.-P. (2019) *Slow decay of impact in equity markets: insights from the ANcerno database.* Market Microstructure and Liquidity, DOI: 10.1142/S2382626619500060.
- Bucci, F., Lillo, F., Bouchaud, J.-P., Benzaquen, M. (2019) *Are trading invariants really invariant? Trading costs matter.* Submitted to Quantitative Finance.
- Bucci, F., Mastromatteo, I., Bouchaud, J.-P., Benzaquen, M. (2019) *Impact is not just volatility.* Quantitative Finance, 19:11, 1763-1766.

Acknowledgements

I want to thank Fabrizio Lillo and Jean-Philippe Bouchaud who supervised me during my doctoral research. I am very grateful for the chance to work with you, for your guidance and for the opportunity to do research in Capital Fund Management. I thank also Stefano Marmi and the Scuola Normale Superiore of Pisa for allowing my visiting period in Capital Fund Management. This research activity was supported by Unicredit S.p.a. under the project "Dynamics and Information Research Institute - Quantum Information (Teoria dell'Informazione), Quantum Technologies".

Chapter 1

Outline of the thesis

This thesis presents a selection of studies on the market impact of *metaorders* - a large orders splitted and executed incrementally in the same direction by the same investor - providing empirical evidences and theoretical models from several viewpoints. In particular, we explore the market impact taking into consideration that market participants have a wide spectrum of reaction timescales and interact reciprocally trading contemporaneously the same asset. Our data-driven analysis is based on a large dataset of metaorders issued by an heterogeneous set of institutional investors in the U.S. equity market and provided by ANcerno Ltd. (formerly the Abel Noser Corporation). The large number and the heterogeneity of the metaorders traded by several financial institutions allows precise measurements of market impact in different conditions and with a reduced uncertainty.

Our results are interesting from two rather different points of views. One is that they represent a significant improvement in the understanding of the determinants of market impact which is the main component of trading costs for institutional investors and at the same time it constitutes an important aspect for the stability of financial markets. The second aspect is that we are entering an era where the availability of large datasets allow more and more to test accurately theories concerning economical and financial issues with standards comparable to those of natural sciences.

In this chapter we summarize the main contributions presented in Chapters 4 through 7 and relate them to the existing literature.

1.1 Market impact of metaorders

It is commonly acknowledged fact that market price moves during the execution of an order - in average it goes up for a buy trade and it goes down for a sell order. This phenomena, known as market impact[1], is crucial in studying the price formation process as well as the optimal execution and transaction cost analysis related problems. Indeed, in order to know whether a trading strategy will be profitable, it is essential to monitor

[1] It is also known as price impact.

transaction costs linked to the market impact. Although the subject importance, there are few research articles pertaining to the empirical estimation of market impact for metaorders, mostly due to the fact that metaorder information is not publicly available. In fact, metaorders have started being recorded in a systematic way only recently and the majority of the database are proprietary, then not readily accessible to academic researchers.

Nevertheless, at the heart of all the empirical and theoretical studies of metaorder impact lies a very simple question: How does the market impact depend on the size Q of the executed metaorder? To answer this question the market impact for a metaorder with Q shares executed in a time interval $[t_\mathrm{s}, t_\mathrm{e}]$ is quantified by the conditional average

$$I(Q) = \mathbb{E}[\epsilon \cdot (s(t_\mathrm{e}) - s(t_\mathrm{s}))|Q] \tag{1.1}$$

where $\epsilon = \pm 1$ is the sign (buy/sell) and $s(t)$ is the rescaled price given by the logarithm of the average market price $S(t)$ at time t normalized by the daily volatility σ_d, i.e. $s(t) = \log(S(t))/\sigma_\mathrm{d}$. Although it may appear intuitive, many models in the theoretical economics predict a market impact $I(Q)$ linear in function of Q [1]. However, there is a growing empirical evidence that the market impact is concave and well described by a *square-root law*, i.e.

$$I(Q) = Y \times \left(\frac{Q}{V_\mathrm{d}}\right)^{\delta} \tag{1.2}$$

with δ an exponent in the interval $0.4 - 0.7$, Y a numerical prefactor of order one called Y*-ratio*, and V_d the total daily traded volume [2, 4, 5, 9, 12, 46, 50, 82]. Note that the square-root law Q depends on the volume fraction $\phi = Q/V_\mathrm{d}$ and it results to be surprisingly universal across different financial products (equities, futures, bitcoin, and options), time periods, market microstructure (small ticks vs. large ticks), underlying trading strategies, and execution styles.

The non-additivitiy of the square-root law implies a natural question concerning the *interaction* between metaorders executed simultaneously on the same asset. More precisely, one may wonder whether the simultaneous impact of several metaorders could substantially alter the square-root law or conversely whether the square-root law might itself result from the interaction of different metaorders. To the best of our knowledge this is an open question since earlier empirical analyses are mostly based on proprietary data from single financial institutions. It follows that there is little insight in literature about *interaction* effects between metaorders simultaneously executed by several market participants on the same asset. In fact, financial markets are the arena of a collective hide-and-seek game between heterogeneous market agents which are trading over a broad range of timescales and for multiple purposes. For this reason in this thesis we present several studies which for the first time break down market impact taking into account *multi-timescales* and *multi-agents* interaction effects. The results discussed confirm the validity of the square-root law but put also in evidence the possible and relevant deviations from it as a consequence of these *interactions*.

1.1.1 Effect of multiple timescales on market impact

In the recent literature, so called *latent limit order book* (LLOB) models [2, 7, 82] have proven to be a fruitful framework to theoretically address the question of the market impact of metaorders. The latent limit order book is an ideal and not measurable order book where all the traders intentions - the latent and the visible ones - are considered: this hypothesis is supported by market data, which put in evidence that only a very small fraction of the daily volume traded on the market is instantly available in the real order book [2]. In fact, the vast majority of the daily traded volume progressively reveals itself as trading proceeds: liquidity is then a dynamical process which takes into account that traders tend to hide their intentions as long as they can, since they have no incentive in giving away private information too soon by adding orders to the real limit order book.

In the LLOB models the liquidity dynamics is based on a continuous *mean field* setting where each agent acts randomly and independently from all the others: each trader can then deposit new orders, remove the old ones and change their mind with a diffusive term. This is formally described by a set of reaction-diffusion equations where assuming finite cancellation and deposition rates allows to capture the multi-timescales of the liquidity dynamics [7, 69]. In fact, in view of the way financial markets operate it is natural to consider agents displaying a broad spectrum of timescales, from low frequency institutional investors (*slow* agents) to high frequency traders (*fast* agents). In this multi-timescales set-up it is then predicted a crossover between a linear market impact regime and a square-root regime in function of the metaorder volume: the high frequency liquidity dominates the total market activity while its low frequency counterpart contributes to shape the concavity of the market impact.

Our contribution. In Chapter 4 we test empirically for the first time the multi-timescales LLOB dynamical theory of liquidity which makes specific predictions about the shape of the market impact: a crossover from a linear in volume behaviour for small volumes to a square-root behaviour for large volumes [7, 69]. Allowing at least two characteristic timescales for the liquidity (*fast* and *slow*) we find that the data supports this crossover described by

$$I(Q) \propto \sqrt{Q}\,\mathcal{F}(\eta) \tag{1.3}$$

with η the participation rate - the ratio between the quantity traded Q and the volume traded by the market during the execution duration $T = t_\mathrm{e} - t_\mathrm{s}$ - and $\mathcal{F}(\eta)$ a scaling function given by

$$\mathcal{F}(\eta) \approx \begin{cases} \sqrt{\eta/\pi} & \text{for } \eta \ll \eta^\star \quad \text{"small participation rate regime"} \\ c & \text{for } \eta \gg \eta^\star \quad \text{"large participation rate regime"} \end{cases} \tag{1.4}$$

with $c = 0.4$ and $\eta^\star \approx 3.15 \times 10^{-3}$ (see Figure 1.1 and refer to the Chapter 4 for details). From Eq. (1.3) it follows that the market impact $I(Q)$ is linear in Q for small Q at fixed T, and crosses over for intermediate Q to a square root regime $\sqrt{Q}$ *independently* from the execution duration T. Whereas a linear regime for small Q was already reported by

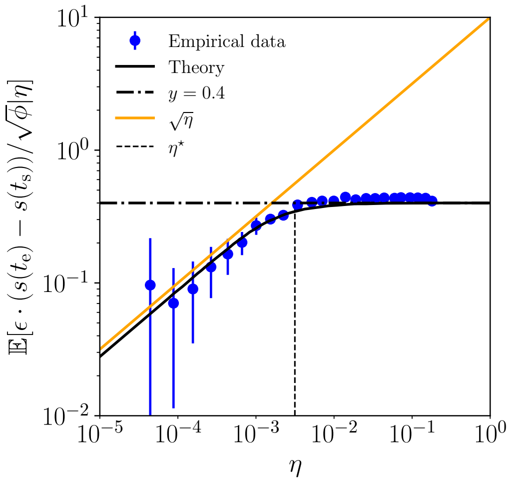

Figure 1.1: Empirical scaling function $\mathcal{F}(\eta)$ estimated by dividing the data into evenly populated bins of constant participation rate η and computing the conditional expectation of $\epsilon(s(t_e) - s(t_s))/\sqrt{\phi}$ for each bin in participation rate, i.e. $\mathcal{F}(\eta) = \mathbb{E}[\epsilon(s(t_e) - s(t_s))/\sqrt{\phi}|\eta]$. The data (blue points) interpolates between a $\sqrt{\eta}$ behaviour observed at small participation rates and an asymptotically constant regime ≈ 0.4 for large η, i.e. for $\eta \gtrsim \eta^\star$ with $\eta^\star \approx 3.15 \times 10^{-3}$, in agreement with the prediction (black solid line) of a multi-timescales LLOB model with *fast* and *slow* agents (see Section 4.6 for details).

Zarinelli et al. [10], the scaling analysis provided by Eq. (1.3) has not been attempted before. In fact, deviations from a pure square-root were observed in [10] where the authors fitted the data with a logarithmic function $\log(a + b\,Q)$, which indeed behaves linearly for small arguments. Furthermore, the fact that market impact in the square-root regime chiefly depends on Q but not on T is compatible with the results of [10], but contradicts many earlier theories that assign the $\sqrt{Q}$ dependence to the *duration* T of the metaorder, as discussed for example in [14, 36, 95]. This point is discussed in detail in Appendix A where we argue that market impact should not be misconstrued as volatility: in particular, the square root law has nothing to do with price diffusion, i.e. that typical price changes grow as the square root of T. We rationalise empirical findings on market impact and volatility by introducing a simple scaling argument in agreement with data.

1.1.2 Multi-agents interaction on market impact

Metaorder information is not publicly available, and earlier analyses mostly based on proprietary data from single financial institutions give little insight about the simultaneous execution of metaorders on the same asset from different investors, which we call *co-impact*. Indeed, even if investors individually decide about their metaorders, they might do so based on the same trading signal. Prices can thus be affected by multi-agents interaction effects such as *crowding* and a natural question comes out: What is the right way to model the total market impact of simultaneous metaorders executed on the same asset and on the same day?

Our contribution. In order to answer to this question we present in Chapter 5 one of the first studies breaking down market impact of metaorders executed on the same asset by different investors (co-impact), and taking into account *crowding* effects. We discuss the limits of the validity of the square-root law on the daily level finding that the market impact of simultaneous daily metaorders is proportional to the square root of their net order flow. Although this is in agreement with the intuition that the market does not distinguish the different individual metaorders, we also found that both the number of investors simultaneously trading on a stock and the *crowdedness* of their trade (measured by the correlation between their metaorder signs) are important factors determining the market impact of a given metaorder.

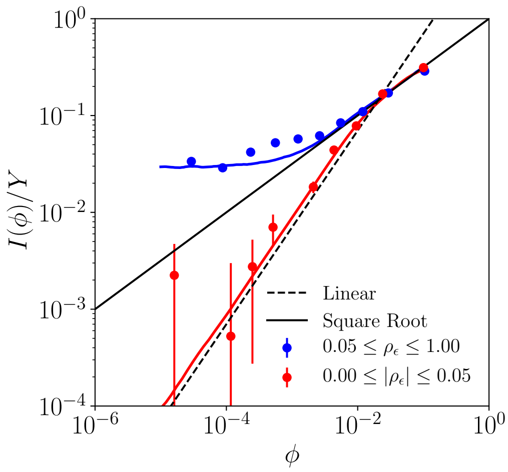

Figure 1.2: Empirical evidence of crowding effects on market impact: comparison between calibrated sign-correlated models (colored lines) and empirical data (colored points) splitted in two samples respectively with crowded (blue points) and not crowded (red points) metaorders. The crowdedness is measured through the sign-correlation estimator ρ_ϵ and the theoretical curves are calculated through numerical simulations as explained in Section 5.5: as evident in figure the calibrated sign-correlated models reproduce well the deviations from the square-root law (black solid line) with both a linear regime and a constant price impact when the order size $\phi \to 0$.

In such a way to have insights on co-impact let us consider the following example: imagine that simultaneously to the considered buy metaorder (with volume fraction $\phi > 0$), another metaorder with the same sign and volume fraction $\phi_m > 0$ is also traded. Since the square-root law applies for the combined metaorders (as confirmed from data), the observed impact should read

$$I(\phi + \phi_m) \propto \sqrt{\phi + \phi_m}. \tag{1.5}$$

It follows that the observed market impact tends to an intercept value $\sqrt{\phi_m}$ when $\phi \to 0$, behaves linearly when $\phi \ll \phi_m$ and as a square root $\sqrt{\phi}$ when $\phi \gg \phi_m$. From this simple but pragmatic example we then based the construction of a theoretical framework to understand when a single investor will observe a square-root impact, and when *crowding effects* will lead to deviations from such a behavior. In fact, considering in the models the sign-correlation between metaorders we are able to reproduce very well the different regimes of the empirical market impact curves in function of the volume fraction ϕ as evident from Figure 1.2. It emerges that any intercept in the empirical market impact curve can be interpreted as a *non zero* correlation with the rest of the market and therefore as a crowdedness metric. Conversely, when the number of metaorders is sufficiently large and the investor is not crowded with the market, then a pure square-root law is recovered.

1.2 Price impact relaxation

After understanding the market impact of a metaorder, another natural question comes out: What happen to the price dynamics when a metaorder is completely executed? This question is still a matter of debate, although on general grounds, one expects intuitively that the price impact starts to relax from its peak value (quantified by the square-root law). However, from the empirical data it emerges that the order flow associated to the execution of a metaorder tends to be autocorrelated in time since the same trading decision to buy or sell might still be valid on the next day, week or month. It follows that the price dynamics seen after completion of a metaorder can be characterized by an apparent plateau[2] reflecting both its own decaying impact and the impact dynamics of correlated metaorders as well.

In this regard, Farmer et al. [15] argue from fair pricing and no arbitrage arguments that the asymptotic value of the price impact should be a fraction of the peak value, in such a way that the average price paid by the buyer or seller is equal to the long-term value. For a square-root market impact it is predicted that the permanent impact is equal to 2/3 of the peak value.

However, as far as empirical data is concerned, the situation is rather confusing, also because the determination of the time when the relaxation terminates is not unique. Some papers, determining permanent impact shortly after the end of the metaorder, report results compatible with the 2/3 value predicted by Farmer et al. theory [5, 8, 10, 11, 88], although Gomes et al. [11] describe a more complex picture, where *informed* and *unformed* orders lead to a very different impact relaxation pattern. In the former case, the impact seems to relax towards the predicted 2/3 value, while in the latter case, price impact appears to relax all the way to zero. Brokmann et al. [12], on the other hand,

[2]It is also defined permanent impact.

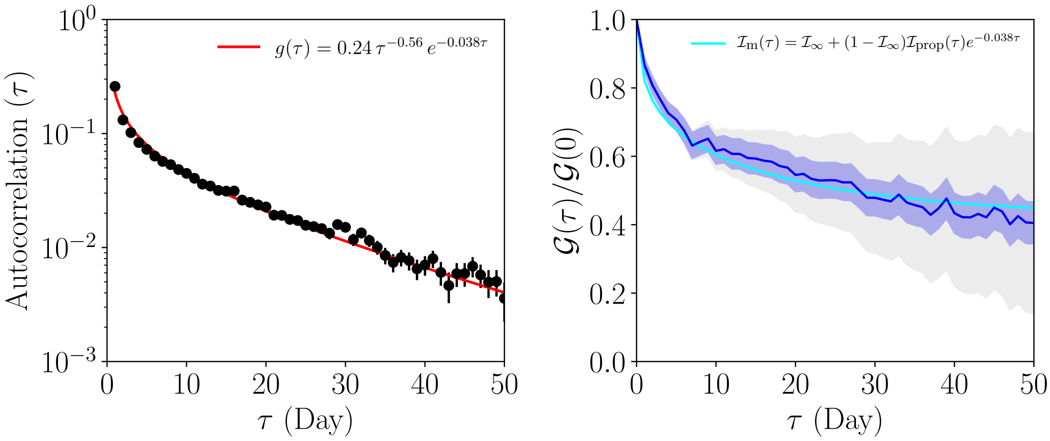

Figure 1.3: (Left panel) Empirical autocorrelation of the net daily order flow imbalance as a function of the lag τ (measured in days) and computed averaging over all stocks in the ANcerno dataset. This autocorrelation persists over many days, in agreement with the fact that the same trading decision to buy or sell might still be valid on the next day, week or month: we fit it with an exponentially truncated power law $g(\tau) = a\tau^{-\gamma}e^{-b\tau}$ with $a = 0.24 \pm 0.04$, $b = 0.038 \pm 0.002$ (corresponding to $1/b \simeq 26$ days), and fixing $\gamma = 0.56$ as discussed in Section 6.5. (Right panel) Price impact relaxation $\mathcal{G}(\tau)/\mathcal{G}(0)$ over multiple days (blue solid line) estimated from the propagator kernel $\mathcal{G}(\tau)$ defined by the deconvolution method described in Section 6.5. The fit of $\mathcal{G}(\tau)/\mathcal{G}(0)$ is represented by the cyan solid line and corresponds to the exponentially truncated modified propagator model $\mathcal{I}_{\rm m}(\tau) = \mathcal{I}_\infty + (1 - \mathcal{I}_\infty)\mathcal{I}_{\rm prop}(\tau)e^{-b\tau}$ with $b = 0.038$ (fixed in agreement with the left panel), an asymptotic decay level $\mathcal{I}_\infty \approx 0.42 \pm 0.01$ and the propagator model $\mathcal{I}_{\rm prop}(\tau) = (1+\tau)^{1-\beta} - \tau^{1-\beta}$ where $\beta = (1-\gamma)/2 = 0.22$. The error bars on the graph are (i) bootstrap errors (blue region) and (ii) cumulated regression errors (grey region) (see Section 6.5 for details).

underline the importance of metaorders split over many successive days, as this may strongly bias upwards the apparent plateau value. After accounting for the metaorder autocorrelation, Brokmann et al. [12] conclude that the price impact decays as a power-law over several days, with no clear asymptotic value, in agreement also with the work of Bacry et al. [9].

Our contribution. In Chapter 6 we revisit the price impact relaxation issue both at the intraday as well as at multiple days level using the ANcerno database. We find that the price relaxation takes place as soon as the metaorder ends and it continues in the following days with no apparent saturation at the plateau value predicted by the Farmer et al. theory [15]. Furthermore, we note that the overnight contribution to the impact decay is negligible in agreement with the idea that it takes place in volume time rather than in physical time. However, due to a significant autocorrelation between the daily net order flow (see left panel in Figure 1.3), a careful deconvolution of the observed impact must be performed in such a way to describe the price impact decay over multiple days (see Section 6.5 for details). It follows that the price impact relaxation is described by an exponentially truncated modified propagator model (see right panel in Figure 1.3) with a power-law behavior at short time scales and a non-zero asymptotic value $\mathcal{I}_\infty$ at long time scales ($\sim$ 50 days) equal to $\approx 1/3$ of the peak impact. Our results match qualitatively those of Brokmann et al. [12] obtained on a smaller set of proprietary metaorders executed by Capital Fund Management (CFM).

1.3 Trading invariance principle

Understanding the dynamics of financial markets is of obvious importance for the financial industry, but also for decision makers, central bankers, and regulators. It is also a formidable intellectual challenge that has attracted the interest of Benoit Mandelbrot who was the first to introduce the idea of scaling [64], a concept that in fact blossomed in statistical physics before getting acceptance in economics and finance (for a review, see [65]): scaling laws are relations between quantities, in which they typically appear in terms of power. In the last twenty years, many interesting scaling laws have been reported, from the square-root law to others concerning different aspects of price and volatility dynamics: one particular question that has been the focus of many studies is the relation between volatility and trading activity, measured as the number of trades and/or the volume traded [59, 112–116].

Revisiting these results, Kyle and Obizhaeva (KO) recently proposed an inspiring hypothesis, coined as the *trading invariance principle* [51, 52]. This principle supports the existence of a universal invariant quantity I_{ko} expressed in dollars, independent of the asset and constant over time, which represents the average cost of a *bet*, i.e. a trading idea typically executed in the market as a sequence of many trades over several days. More in detail, the KO trading invariance principle predict that the quantity $I_{\mathrm{ko}} = \mathcal{W}/N^{3/2}$, where $\mathcal{W}$ is the exchanged risk (volatility $\times$ volume $\times$ price) and N is the number of bets, is invariant. This invariance implies the scaling $\mathcal{W} \sim N^{3/2}$ coined as 3/2-law and which can be interpreted with different degree of universality: *no universality* (the 3/2-law holds for some financial instruments only), *weak universality* (the 3/2-law holds but with a non-universal value of I_{ko}), and *strong universality* (the 3/2-law holds and I_{ko} is constant across assets and time) [57].

In the recent literature the KO trading invariance principle has been empirically investigated at single-trade level rather than at the bet level mainly as a consequence of the fact that identifying a bet in the market is not a straightforward task[3]. Although empirical data at single-trade level revealed that while the 3/2-law is very robust, it emerges that the I_{ko} is *not* invariant as it is asset and time dependent [51, 55–58]. However, since single transactions are typically not the same as single bets, to the best of

[3]In [52] Kyle and Obizhaeva tackled this problem exploiting a proprietary dataset of portfolio transitions.

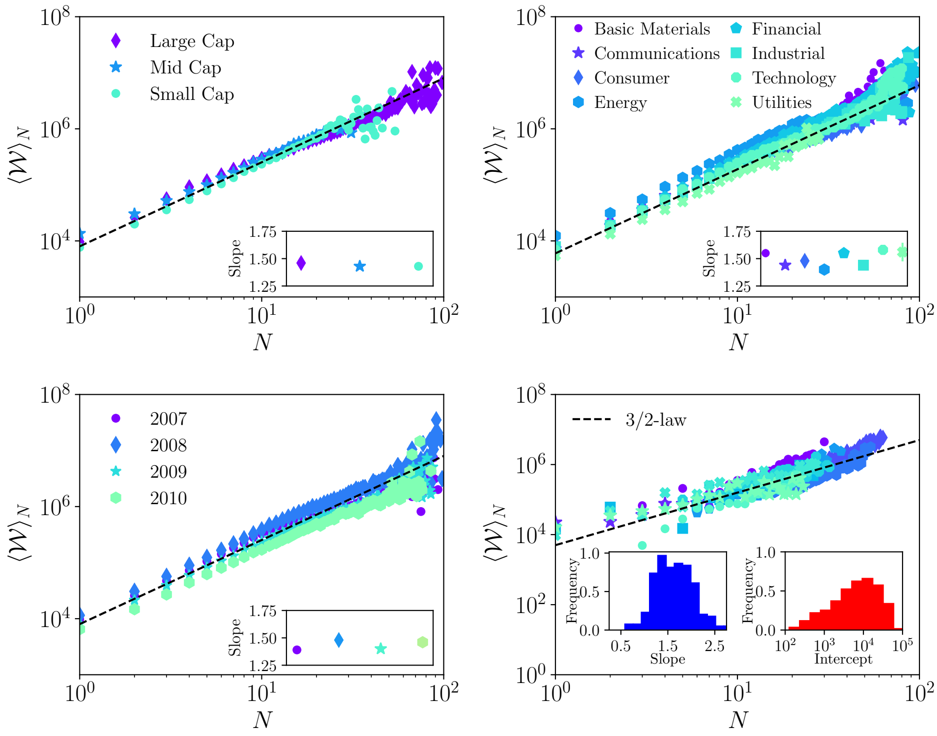

Figure 1.4: Empirical evidence of the scaling 3/2-law from the mean daily exchanged risk $\langle \mathcal{W} \rangle_N$ conditional on the daily number N of metaorders per asset for different market capitalisations (top left panel), economic sectors (top right panel), and time periods (bottom left panel). The insets show the slopes obtained from linear regression of the data, firstly averaged with respect to N and secondly log-transformed. The bottom right panel shows a plot of $\langle \mathcal{W} \rangle_N$ as function of N for a subset of 10 stocks chosen randomly from a pool of around three thousand U.S. stocks: the two insets represent respectively the empirical distribution of the slopes and of the y-intercept obtained from linear regression of a larger sub-sample of 200 stocks randomly chosen, firstly averaged with respect to N and secondly log-transformed considering each stock separately. Although the 3/2-law works well it is evident that I_{ko} is asset dependent and then not invariant (see Section 7.3 for details).

our knowledge the testing of the KO trading invariance principle at the bet level is still missing. This is in fact the aim of the Chapter 7.

Our contribution. In Chapter 7 we revisit the KO trading invariance hypothesis [51,52] by empirically investigating the large dataset of metaorders provided by ANcerno. As suggested by Kyle and Obizhaeva metaorders can be considered a proxy of bets. Focusing at the daily level we find that the 3/2-law between $\mathcal{W}$ and N works surprisingly well and it is robust against changes of year, market capitalisation, economical sector, and assets as evident from the several panels in Figure 1.4. However, our empirical analysis clearly shows that the quantity I_{ko} is not invariant for a wide range of assets (see bottom right panel in Figure 1.4) and we argue in favour of a weak universality degree. Since a very high correlation between I_{ko} and the average total trading cost $\mathcal{C}$ (spread and market impact) per metaorder is measured we then propose a new invariant defined as the ratio $\mathcal{I} = I_{\text{ko}}/\mathcal{C}$ finding a large decrease in variance as evident from Figure 1.5. Finally, we exhibit the microstructural origin of the small dispersion of the new invariant $\mathcal{I}$ which is mainly driven by (i) the scaling of the spread with the volatility per transaction, (ii) the near invariance of the distribution of metaorder size and of the volume and number fractions of metaorders across stocks.

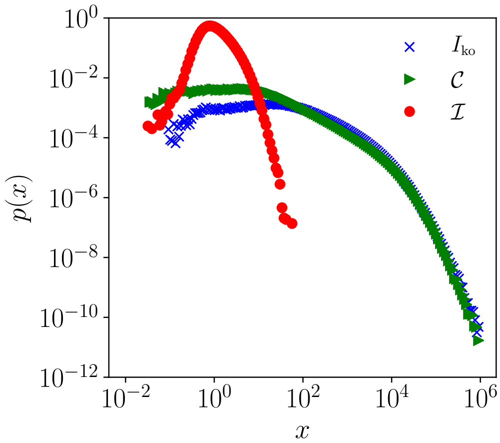

Figure 1.5: Empirical distributions in log-log scale of the KO invariant $I_{\text{ko}} = \mathcal{W}/N^{3/2}$, of the daily average metaorder total trading cost $\mathcal{C}$, and of the dimensionless invariant $\mathcal{I} = I_{\text{ko}}/\mathcal{C}$: note that renormalizing the KO invariant I_{ko} by the average metaorder total trading cost $\mathcal{C}$ implies a large decrease in variance as evident in figure. For details on the computation of the involved observables see Section 7.4.

Chapter 2

Introduction

In the last years, the availability of high-quality data have considerably revolutioned the investigation of financial markets. The information that can be extracted from these data allows to have new insights on how financial markets works and it is at the basis of the market microstructure field. As the name suggests, the market microstructure concerns the details of how specific trading mechanisms affect the price formation for financial securities providing insight into the emergence of complex phenomena that have been widely reported but poorly understood in the financial literature [28–30]. In this view, the understanding of this phenomenology is of obvious importance for many practical purposes, for instance, the quantifying of market impact[1], the reduction of execution costs, the design of trading strategies, the organisation of markets and the lowering of financial risks [32, 34].

As this growing data-driven research strand comprises many diverse subfields, uniting researchers from various disciplines including economics, mathematics, physics, econometrics and data science as well as financial practitioners and regulators, a comprehensive overview is beyond the scope of this chapter. Instead we provide a comprehensible introduction to the topics that are most relevant to this thesis.

2.1 Limit order book

Financial markets allow different sources of information to be processed and transformed into a single number: the price. They are complex systems where many agents, called traders, act with the purpose of maximizing their profits. In modern electronic markets, it is common that traders interact through a limit order book (LOB) in a continuous double auction. Continuous is referred to time, meaning that at any moment market participants can take an action on the limit order book, while double auction refers to the fact that the limit order book is divided into two sides: the buyers on the bid side and the sellers on the ask side, as shown in Figure 2.1. The highest buy order is the best bid while the lowest sell order is the best ask and their difference is called the bid-ask

[1] It is also known as price impact.

spread. The mid-price is defined as the mid-point between the best bid and the best ask and the discretization of the price axis is quantified by the tick size.

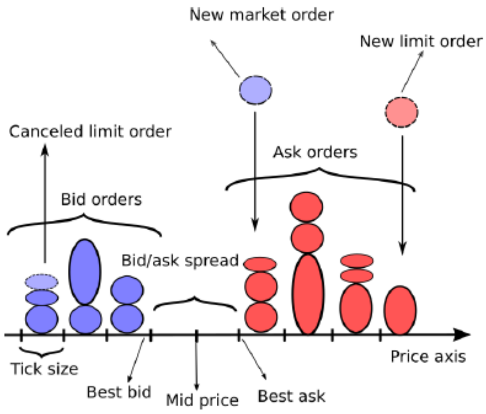

Figure 2.1: A sketch of the limit order book - in blue the bid side and in red the ask side - with the arrows indicating the main quantities defined in the main text. For example the figure shows a canceled limit order on the bid side and a market and limit orders on the ask side.

From the operational point of view, each trader can post an order that will be characterized by three key quantities: the sign (buy or sell), the number of shares, and the price at which to trade it. The actions that can be done in the limit order book are of three types:

- Market order which implies an immediate transaction at the best available opposite price.
- Limit order which allows the trader to secure the price at which its order will be executed, but not the time.
- Cancellation order which allows to change a previous taken position inside the limit order book removing the order.

All these three types of orders are very important to describe the dynamics of the limit order book [32, 34] and obviously the decision whether to post a market or a limit order is mainly dictated by the need that the trader has to have the deal done. In fact bid orders show market interest in buying while the ask orders show market interest for participants who are willing to sell but necessarily at a price higher than the best bid. If a participant wants to buy (sell) now, then they need to hit the ask (bid) and to pay a high (low) price. If a participant wants to buy (sell) but is more patient, they may choose to join those on the bid (ask) and to wait for someone to hit them at a low (high), and therefore a better price. It follows that the dynamics of the limit order book is the result of a fine tuning between the behavior of liquidity providers and liquidity takers.

Historically, the task of supplying liquidity by permanently maintaining limit orders in the LOB was assumed by designated market-makers who, in exchange of this service, kept the spread: they offered to buy at a price lower than their sell price, leading to a profit on each transaction with an unchanged mid-price. All other actors are forced to interact with a market-maker were liquidity takers. In reality, the idea of market-maker profiting on each transaction is limited by the challenge represented by the *adverse selection*: if an informed trader has an accurate prediction about the future price of an asset initially not available to the market-makers, then he can profit by entering a transaction with them. To reduce this information mismatch, the market-maker seeks to process any new piece of public information as soon as possible re-adjusting accordingly the offered quotes. In nowadays electronic trading systems, anyone can be a market-maker with its own strategy: they are usually high frequency traders who react in a much faster way with respect to the low frequency traders creating a multi-time scales liquidity in the limit order book [34].

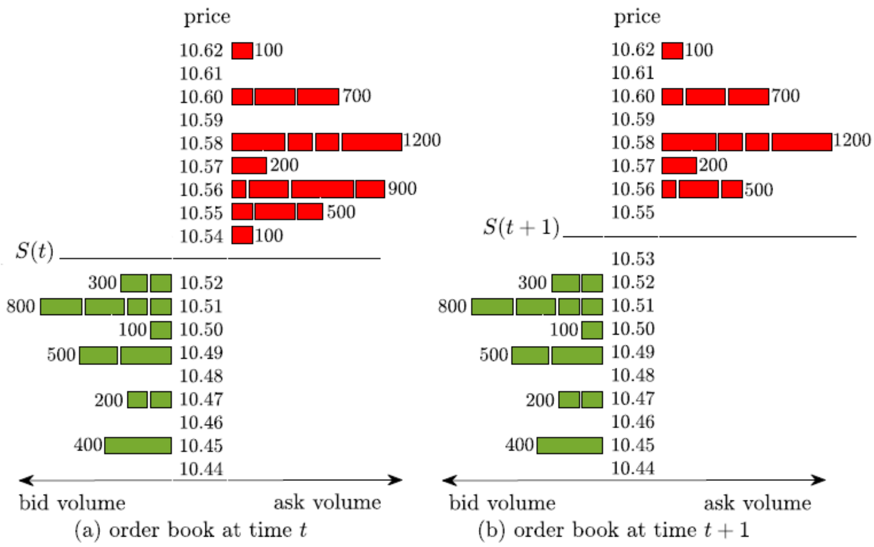

Figure 2.2: Snapshot of the limit order book before and after the execution of a buy market order with a volume of 1000 shares. At each price level, a bar stands for a limit order and its width represents the volume that will be traded. To note that the mid point price $S(t) \to S(t+1)$ moves up after the execution of a market order.

2.1.1 A concrete example

Let us take a concrete example to illustrate how the continuous double auction works. Figure 2.2 shows a snapshot of the limit order book before and after the execution of a buy market order. As shown in the figure, there are 100 shares available at the best ask at time t. If the market order is only for buying 100 shares, it will be fully executed with a trade price of 10.54 dollars consuming the volume at the best ask. However, if the market order is for buying 1000 shares, the case will be different. In particular, the market order will be executed partly at the price of 10.54 dollars for 100 shares. To fulfill the demand of 1000 shares, the market order will continue being executed at the price of 10.55 dollars for 500 shares and at the price of 10.56 dollars for the remaining 400 shares. Consequently, this large market order changes the ask price from 10.54 dollars at time t to 10.56 dollars at time $t+1$. Such a price change due to a trade is termed as the price impact and it is an all-too familiar issue for traders who need to buy or to sell large quantities of an asset since the impact of their earlier trades makes on average the price of their subsequent trades worse (see Section 2.3 for more details).

The previous example illustrates the extra cost which can incur during a buy (sell) order due to the scarcity of supply (demand). Although this effect is often insignificant for small trades, it becomes relevant when a trader wishes to execute large volumes: the impact of their orders is noticeable and must be taken into account as an additional cost. It has been measured that the instantaneous volume displayed on the limit order book is approximately 0.1% of the total daily traded volume [2] and this is a coherent with the fact that market participants want to release as little information as possible about their intentions, at least until they have a fair confidence that their orders will be executed in a reasonable amount of time. This implies that when a trader seeks to execute a large transaction it must be sliced and diced in such a way to be executed incrementally.

2.2 Slice and dice: the origin of a metaorder

It is common practice that a large portion of transactions are executed by algorithms on electronic trading platforms. This shift from floor trading to largely automated trading on electronic markets has led to an increase in frequency and volume of transactions. Most large market participants, such as investment banks, hedge funds, and proprietary trading firms develop sophisticated algorithms to take advantage of the possible arbitrage opportunities and to profit from them due to the large volume of trades that they make. In this view, these market participants are faced with several high-level tasks such as to make trading strategy decisions as for example which asset to trade and when, whether to buy or to sell, at which price and so on.

When an asset manager takes the decision to buy or to sell some quantity Q of a given asset it would be ideal to buy or to sell it immediately at the best available price. However, for the range of volume typically executed by large financial institutions, there is rarely enough available volume at the best price to absorb the required quantity at all once. For this main reason, it is usual to slice and dice the full quantity Q into smaller pieces, called *child orders*, which are executed sequentially using both market and limit orders over a period which might spans several minutes to several days. This sequence of transactions executed in the same direction (buy/sell) and belonging to the same trading decision is called a *metaorder*. In practice, the trading activity associated to the execution of a metaorder is articulated in two parts:

- At the investment-decision stage the trader determines the direction (buy/sell), the number of shares Q and the time duration T over which to execute the metaorder. These decisions are usually based on some belief of the asset manager about the future price of the asset.

- The execution stage, during which trades are conducted to obtain the required quantity at the best possible price within the prefixed time window T. This step is sometimes delegated to a broker, who seeks to achieve specified execution targets by performing incremental execution of the metaorder in small chunks (see Appendix. 2.A for an introduction to some common trading execution algorithms).

In this way investors try to minimize their information leakage to the other market participants and at the same time to manage the execution costs induced by the market impact. However, note that to find statistical regularities for the execution costs of a metaorder require to think in statistical terms. In fact, breaking up a metaorder and executing it sequentially means that each scenario is different and the empirical derivation of an operational formula for the market impact is an exercise of averaging over many metaorders. This remark is at the basis of the formal market impact definition introduced in the following section.

2.3 Market impact and the square-root law

On average buy trades push up the price and vice versa sell trades push it down. This statistical effect is quantified by the market impact which describes how much price dynamics is influenced by the order flow of a metaorder.

More precisely, for a metaorder with Q shares executed in a time interval $[t_{\mathrm{s}}, t_{\mathrm{e}}]$ the market impact is quantified by the conditional average

$$I(Q) := \mathbb{E}[\epsilon \cdot (s(t_{\mathrm{e}}) - s(t_{\mathrm{s}}))|Q] \tag{2.1}$$

where $\epsilon = \pm 1$ is the sign (buy/sell) and $s(t)$ is the rescaled price given by the logarithm of the average market price $S(t)$ at time t normalized by the daily volatility σ_{d}, i.e. $s(t) := \log(S(t))/\sigma_{\mathrm{d}}$. Naively, it might seem intuitive that the market impact of a metaorder should scale linearly with its size. This is in fact in agreement with the prediction of the seminal microstructure model proposed by Kyle in 1985 [1]. However, the last decades have witnessed mounting empirical evidence invalidating this linear impact framework (see for example Figure 2.3) and revealing a concave and approximately square root scaling, i.e.

$$I(Q) = Y \times \left(\frac{Q}{V_{\mathrm{d}}}\right)^{\delta} \tag{2.2}$$

with δ an exponent in the interval $0.4 - 0.7$, Y a numerical prefactor of order one called Y-ratio , and V_{d} the total daily traded volume [2, 4, 5, 9, 12, 46, 50, 82]. This non-additive scaling, coined as the *square-root law*, asserts that after a trade of volume $Q/2$, the next half-one $Q/2$ will have less impact on the price change (~ 40 % of the first $Q/2$). Since it has been empirically tested for several markets and instruments as for example

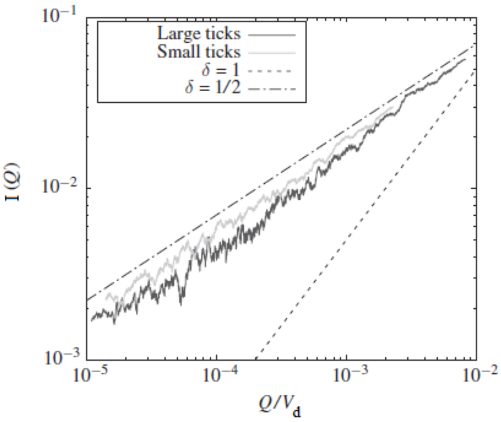

Figure 2.3: Market impact measured using metaorders executed by Capital Fund Management (CFM) on futures market during the time period from 2007 to 2010 (reproduced from Ref. [2]). The market impact $I(Q)$ is represented in function of the ratio Q/V_d on doubly logarithmic axes. The black curve is for large-tick futures and the grey one is for small-tick futures. For comparison, it is also shown a dash-dotted line of slope $\delta = 1/2$ corresponding to a square-root impact and a dotted line with slope $\delta = 1$ corresponding to a linear impact.

stocks [4, 5, 12, 50, 88], future contracts [2], options [13], bitcoin [89], and it seems to not depend on the geographical zone or time period [82] a spree of theoretical research activity started in such a way to understand its origin. In the next section we recall some of these most relevant theoretical frameworks developed in the literature.

2.4 Some theoretical models for market impact

One of the principal motivations of this thesis is to test fundamental theories for market impact that we describe in the following. For more details we refer the reader to the original papers.

2.4.1 Kyle model

In the classical financial literature one of the central model for the market impact is represented by the Kyle model [1] which describes in presence of information asymmetry how this information is incorporated into price. In particular, the theory takes in consideration a single period equilibrium[2] between three types of market participants, each representing a well-identified trading behaviour: an informed trader competes with a market-maker who provides liquidity for every trade in presence of noise traders submitting a random trade volume. At the beginning of the trade period all market par-

[2]In its original work [1] Kyle describes also the multiple periods and the continuous time variants.

ticipants trade in an asset with an initial price S_0 and a final normal liquidation price $S_F \sim \mathcal{N}(S_0, \Sigma_0)$ with mean S_0 and variance Σ_0. From the available private information the informed trader knows exactly the liquidation price S_F that the asset will have at the end of the period and then chooses to buy ($\epsilon = +1$) or to sell ($\epsilon = -1$) a trade volume Q in such a way to optimize her expected profit. At the same time the noise traders submit a normal trade volume $V_{\text{noise}} \sim \mathcal{N}(0, \Sigma_V)$ with mean zero and variance Σ_V. The market maker observes then a total net order flow $\Delta V = \epsilon Q + V_{\text{noise}}$ which is matched with her own inventory through a rule-based clearing price $\hat{S}$. However, knowing that there is an informed trader who wants to trade as much as possible to exploit her informational advantage, the market maker protects herself by setting a clearing price that is increasing in the total net order flow ΔV. In particular, assuming that

1. the market maker, supposed competitive, realizes in average a null profit and
2. the strategy of the informed trader is optimal with respect to the pricing rule of the market maker,

Kyle shows the existence of a unique *market equilibrium* if the pricing rule for the clearing price $\hat{S}$ is linear, i.e.

$$\hat{S} = S_0 + \lambda_{\text{kyle}}(\underbrace{\epsilon Q + V_{\text{noise}}}_{\Delta V}). \tag{2.3}$$

In Eq. (2.3) the Kyle's coefficient $\lambda_{\text{kyle}} = \frac{1}{2}\sqrt{\frac{\Sigma_0}{\Sigma_V^2}}$ measures the market illiquidity, i.e. larger the coefficient, the more a given volume impacts the price and the more expensive is the trading activity. From this linear clearing price rule it follows that the profit maximisation for the informed trader is given by

$$\hat{Q} = \text{argmax}_Q \mathbb{E}[\pi] \tag{2.4}$$

where

$$\pi = \epsilon Q \times (S_F - S_0 - \lambda_{\text{kyle}}(\epsilon Q + V_{\text{noise}})). \tag{2.5}$$

Since the expected value of the random trade noise V_{noise} is zero, one has $\mathbb{E}[\Delta V] = \epsilon Q$. This leads from Eq. (2.4) to a quadratic maximization problem with solution

$$\hat{Q} = \frac{S_F - S_0}{2\lambda_{\text{kyle}}} = (S_F - S_0)\sqrt{\frac{\Sigma_V^2}{\Sigma_0}} \tag{2.6}$$

which implies that in such a way to maximize her profit the informed trader should trade a quantity $\hat{Q}$ proportional to the mispricing $S_F - S_0$ and that is greater when it is possible to hide her demand in the noise traders liquidity (measured by Σ_V). It follows that the informed trader performs on average a profit equal to

$$\mathbb{E}[\pi] = \frac{(S_F - S_0)^2}{2}\sqrt{\frac{\Sigma_V^2}{\Sigma_0}} \tag{2.7}$$

at the expense of the noise traders which are loosing their money.

In summary, the Kyle model shows that the total order flow impacts the price because of its expected information content and in the particular case of Gaussian distributed trading volumes the market impact scales linearly with the informed trader's volume, i.e. $I(Q) \propto Q$. Although this prediction is in clear contrast with the empirical data (as discussed for example in the previous section), the Kyle model is often cited as the foundation of the field of market microstructure and it is considered as a starting point for more realistic models.

2.4.2 Fair pricing theory

In the fair pricing theory of Farmer, Gerig, Lillo & Waelbroeck (FGLW) [15] a martingale hypothesis and a fair pricing condition are combined to derive the relationship between the distribution of the metaorder size and the shape of the market impact. In analogy to the Kyle approach [1] the market ecology is represented by three types of agent: informed traders, daily traders, and market makers. Informed traders are rational and long-term institutional investors who buy or sell orders following a common information signal α drawn from an exogenously distributions $p(\alpha)$. The daily traders follows their private information signal to generate an order flow which is aggregated to the informed one and executed incrementally over time through *equally sized lots* by taking liquidity supplied by profit maximizing and competitive market makers. The main purpose of the fair pricing theory is then to understand the way how order splitting of a metaorder in n slices of equal size κ at times $t = 1, \cdots, n$ affects the shape of the market impact assuming that the beginning and the end of a metaorder can be detected by all the market participants.

To this aim the fair pricing theory of FGLW firstly imposes the martingale condition for the transaction price $\tilde{S}_t$ during the lifetime of the metaorder, i.e.

$$\mathcal{P}_t(\tilde{S}_{t+1} - \tilde{S}_t) + (1 - \mathcal{P}_t)(S_{t+1} - \tilde{S}_t) = 0 \tag{2.8}$$

where $\mathcal{P}_t$ is the probability that the metaorder continues beyond t executions while $(1 - \mathcal{P}_t)$ is the probability that the metaorder stops. The likelihood that the metaorder will persist depends on the order size distribution $p(n)$ and on the number of executions t that are already done. $\tilde{S}_t$ and $\tilde{S}_{t+1}$ represents the prices before and after the execution of the $(t + 1)$-th slice and S_{t+1} is the post-trade price in the case that the metaorder is fully executed in t slices. From the martingale condition expressed in Eq. (2.8) it follows that

$$\frac{\tilde{R}_t}{R_t} = \frac{1 - \mathcal{P}_t}{\mathcal{P}_t} \tag{2.9}$$

where $\tilde{R}_t = \tilde{S}_{t+1} - \tilde{S}_t$ and $R_t = \tilde{S}_t - S_{t+1}$ are respectively the incremental price responses to the continuation and to the completion of the metaorder. However, another condition is required to derive the values of $\tilde{R}_t$ and R_t at each auction t and therefore, to describe the shape of the market impact.

This second condition is given by the fair pricing assumption which states that the post-trade price S_{n+1} is equal to the average execution price

$$S_{n+1} = \frac{1}{n}\sum_{i=1}^{n}\tilde{S}_i. \tag{2.10}$$

At this point and as shown in the original work of FGLW [15] the combination of the martingale hypothesis together to the fair pricing assumption allows to form a system of equations that can be solved respect to $\tilde{R}_t$ and R_t in such a way to describe the market impact in terms of the probabilities $\mathcal{P}_t$. In the special case of a Pareto distribution of order size, i.e.

$$p(n) \sim \frac{n^{-(\gamma+1)}}{\xi(\gamma)} \tag{2.11}$$

where $\xi(\gamma)$ is the Riemann zeta function, the fair pricing theory predicts that:

1. In the limit of large t, the market impact[3] increases asymptotically as

$$I_t \sim \begin{cases} t^{\gamma-1} & \text{for} \quad \gamma \neq 1 \\ \log(t+1) & \text{for} \quad \gamma = 1. \end{cases} \tag{2.12}$$

2. For $\gamma \neq 1$ the permanent impact - the persistence of a shift in the price after the metaorder is fully executed - behaves asymptotically for large time t as

$$\mathcal{I}_\infty \sim \frac{1}{\gamma}t^{\gamma-1}. \tag{2.13}$$

Combining the previous two results it emerges that the ratio of the permanent impact to the peak market impact is equal to

$$\frac{\mathcal{I}_\infty}{I_t} \sim \frac{1}{\gamma}. \tag{2.14}$$

According to [15] there is a considerable evidence that for most equity markets the metaorder size is distributed in the limit of large size as $p(Q > v) \sim v^{-\gamma}$ with $\gamma \approx 3/2$. This implies that the price reversion after the completion of a metaorder should converge to a permanent impact equal to 2/3 of the peak market impact.

In summary, the fair pricing theory of FGLW predicts that if the metaorder size is distributed as a power law with exponent γ, then the market impact function behaves as Q^δ respect to the metaorder's volume Q with $\delta = \gamma - 1$: to note that this corresponds to the empirical square-root law ($\delta = 1/2$) if $\gamma = 3/2$. However, it is found empirically that the relation between the market impact exponent δ and the power-law exponent γ for the distribution of metaorder sizes does not hold universally [82, 89]. Furthermore, the fair pricing theory of FGLW suggests that the market impact decays instantaneously to

[3] The authors also call it the *peak* market impact.

its asymptotic permanent value after the metaorder conclusion, whereas empirical data reveals that this behaviour is not routinely observed in real markets. Instead, market impact undergoes an initially steep decay, then relaxes very slowly over a period that can span several days to an asymptotic permanent value which is still matter of debate (see Chapter 6 for more details on this topic).

2.4.3 Latent limit order book approach

The limit order book is at the heart of the financial markets behaviour. However, since most of the trading activity is not declared until the very end without appearing inside the LOB, Tóth et al. [2] introduced the idea of the latent limit order book (LLOB) in such a way to take into account all the trading intentions, the latent and the visible ones. In a pragmatic way, such latent limit order book can be seen as a proxy for the supply and demand at the intra-day time scale. In fact, Donier et al. [7] proposed a LLOB model based on a set of reaction-diffusion equations through which the dynamics of the latent bid and ask volume densities are described. It is then shown that under quite general conditions the shape of the latent order book becomes exactly linear around the mid-price and the market impact is square-root on the volume of the metaorder, i.e. $I(Q) \propto \sqrt{Q}$ (see Chapter 4 for more details on this approach).

2.5 Transaction costs

The relevance of the market impact modelling is strictly related to the one of the Transaction Cost Analysis which has become a central issue in the financial industry. In fact, transaction costs are widely recognized as a large determinant for the invesment performance. They not only affect the realized results of an active investment strategy, but they also control how rapidly assets can be converted into cash if it should be the case. Such costs generally fall into two categories:

- Direct cost are represented by the commissions and brokerage fees which are explicitly stated and easily measured. Although they are important and should be minimized, they are not the focus of this thesis.
- Indirect costs are the ones that are not explicitly stated and are related to the market microstructure. For example, these costs are represented by the bid-ask spread and by the impact costs. For large trades, the most important component of these costs is the impact of the trader's own actions on the market which is a consequence of the finite liquidity in financial market. Given its statistical nature it only appears after a careful averaging since otherwise it is invisible to the naked eye and it actually represents the lion's share for transaction cost analysis.

To have an order of magnitude let us assume that an asset manager trades 1% of the daily traded volume V_{d} of a stock characterized by a 2% daily volatility σ_{d}. From the square-root law it emerges that the unitary impact cost $\mathcal{C} = 2/3 \times \sigma_{\mathrm{d}} \times Q^{3/2}/V_{\mathrm{d}}$ is about 15 basis point, i.e. one order of magnitude larger than direct and spread costs which are of the order of 1 basis point. From this example it emerges that modeling and understanding impact is a crucial element to take into consideration if we want to design efficient trading strategies in competitive financial markets. When evaluating investment strategies, the analysis of transactions costs is at risk of being overly simplified, leading to potentially erroneous conclusions about a manager's trading acumen and ultimately suboptimal investment allocation decisions.

Appendix

2.A Trading algorithms

In practice it is a difficult task to execute manually and efficiently a large order over a long time period. Therefore, this is usually done through the use of execution algorithms usually built-in-house by asset managers or proposed as a service by brokers. Some common execution schedules used by practitioner are:

- The time-weighted average price (TWAP) benchmark which aims to obtain an average execution price for the metaorder that is as close as possible to the time-weighted average price present in the market during a fixed time period.
- The volume-weighted average price (VWAP) benchmark which aims to achieve an average execution price for the metaorder that is as close as possible to the volume-weighted average price available during a specified period. It implies that the transaction volumes of the metaorder are higher (lower) during the period of high (low) averaged activity.
- The percentage of volume (POV) for which the volumes of the transaction stays in a narrow band with a width of the order of a few percents around a constant chosen as a fixed fraction of the estimated daily volume.

Obviously, it is possible to devise more sophisticated algorithms but in any case all of them slice large orders into small pieces that will be executed sequentially.

Chapter 3

ANcerno dataset

In the present era technology advance has considerably changed the way securities are traded in financial markets. For the vast majority of developed markets, the algorithmic trading plays a relevant role where sophisticated computer programs automatically make trading decisions and handle the order submission. At the same time this modernisation process provides an incredible large amount of data which can be analysed and used for trading research purposes. In fact, institutional investors are careful to record in a systematic way their trades execution and to use them to estimate how much their strategies incur in trading costs implied by market impact. It follows that a better understanding of market impact should lead to a competitive advantage reducing trading costs and consequently improving assets allocation.

However, the empirical investigation of the market impact requires a clear identification of the sequence and time stamping of each child orders belonging to the same trading decision. A priori these informations are not available in public trade data, which are typically anonymized and then useless for an explicit metaorder identification. Finding a representative dataset of institutional metaorders is never an easy task and in general, the solution is represented by a proprietary trading/broker firm data which contains the informations relative to the placement and execution times, price and number of shares for each metaorder. In our case, the research presented in this thesis is based on the exploitation of the large and heterogeneous ANcerno database which is described in this chapter.

3.1 Data and definitions

Our analysis relies on the database made available by ANcerno Ltd. (formerly the Abel Noser Corporation) which is a widely recognised consulting firm that works with institutional investors to monitor their equity trading costs[1]. The database contains trade-level data gathered on metaorder executions from the main investment funds and brokerage firms in the U.S. equity market. Previous academic studies that use ANcerno

[1]See `www.ancerno.com` for more details

data include for example [10, 43–45, 47–49].

We define a metaorder as a series of successive orders performed by a single investor through a single broker within a single day, on a given stock and in a given direction (buy/sell). The structure of the database provides the information necessary for an explicit metaorder identification represented respectively by

- the total number of shares Q,
- the trade direction $\epsilon = \pm 1$ (buy/sell),
- the time of the first placed order t_{s} (placement time) and the corresponding market price $S(t_{\mathrm{s}})$,
- the time of the last trade t_{e} (execution time) and the corresponding market price $S(t_{\mathrm{e}})$.

It follows that each metaorder is characterised by a broker label, a stock symbol, the total number of shares Q and the times at the start t_{s} and at the end t_{e} of its execution with sign $\epsilon = \pm 1$. The main advantage of this dataset is represented by a clear identification of metaorders relative to the trading activity of diversified institutional investors also if the motivations of the transactions are unknown. In fact the dataset is heterogeneous, containing large institutional orders issued for different purposes and it spans several years from 1999 to 2015: in our case we consider a sufficiently large sub-sample limited to the time period from January 2007 to June 2010 for a total of 880 trading days. To remove possibly erroneous data we follow the procedure introduced in [10]:

- **Filter 1**: We select the stocks which belong to the Russell 3000 index discarding metaorders executed on highly illiquid stocks.
- **Filter 2**: We select metaorders ending before 4:01 p.m.
- **Filter 3**: We select metaorders whose duration $T = t_{\mathrm{e}} - t_{\mathrm{s}}$ is longer than 2 mins.
- **Filter 4**: We select metaorders whose participation rate (the ratio between their quantity and the volume traded by the market between t_{s} and t_{e}) is smaller than 30%.

The extracted sample is represented by ~ 8 million of metaorders distributed quite uniformly across time periods, market capitalisations and economical sectors. These filtered metaorders are around the 5% of the total reported market volume independently of the year and of the stock capitalisation[2]. For comparison we report in Table 3.1 the approximate number of metaorders previously used in literature to empirically investigate market impact: it is evident that our sample is more than one order of magnitude larger than the typical size investigated so far, excluding the work done in [10] where the authors used the ANcerno dataset. The exploitation of a large sample is quite important

[2]Without the above filters, this number would rise to about 10%.

Authors	Publication year	Number of metaorders	Institution
Almgren et al. [4]	2005	700,000	Citigroup
Engle et al. [50]	2008	230,000	Morgan Stanley
Moro et al. [5]	2009	150,000	Inferred
Tóth et al. [2]	2011	500,000	CFM
Bershova and Rakhlin [8]	2013	300,000	Alliance Bernstein LP
Waelbroeck and Gomes [11]	2013	130,000	Various
Mastromatteo et al. [82]	2014	1,000,000	CFM
Brokmann et al. [12]	2014	1,600,000	CFM
Bacry et al. [9]	2014	400,000	Képler Chevreux
Zarinelli et al. [10]	2015	7,000,000	ANcerno
Tóth et al. [13]	2018	$450,000$	CFM
Said et al [88]	2018	1,500,000	BNP Paribas

Table 3.1: The table reports the approximate number of metaorders used in previous studies by several authors together with the corresponding trading institution providing the metaorders.

since market impact measures are very noisy and then larger datasets are ideal to reduce statistical uncertainty and to remove spurious bias effects. Moreover the heterogeneity of financial institutions and brokers in the ANcerno dataset guarantees that our results are not specific to a single execution strategy and limited to a single market participant.

In Figure 3.1, the time series of the metaorders executed on MSFT in the time period from June to July 2009 shows that a significant number of metaorders are active in the same day. In most trading days for a given asset the vast majority of metaorders are executed with the same direction (buy/sell). Unfortunately, from the available information in the dataset it is not possible to know the conditions and characteristics of the metaorder execution. For example we do not know if the metaorders were executed for cash reasons or were informed trades as for example for the dataset used in [11]. Similarly, we do not have information relative to the execution trading algorithm used by the broker and if the trading size was conditioned on movement of the price during execution of the metaorder. Anyway we believe that this weakness of the ANcerno dataset does not change the conclusions discussed in the following chapters of this thesis.

3.2 Descriptive statistics

In this section we introduce the main observables used to describe metaorder execution and we discuss some of their summary statistics. To this aim we follow the common practice to measure time in volume unit where time is moved forward according to the volume traded in the market. For a trading day, let $V(t)$ the total volume traded by whole the market from the opening to the physical time t. We define then the volume time as $v(t) := V(t)/V(t_\mathrm{c})$ where t_c is the daily closing time and $V(t_\mathrm{c}) = V_\mathrm{d}$ is the volume

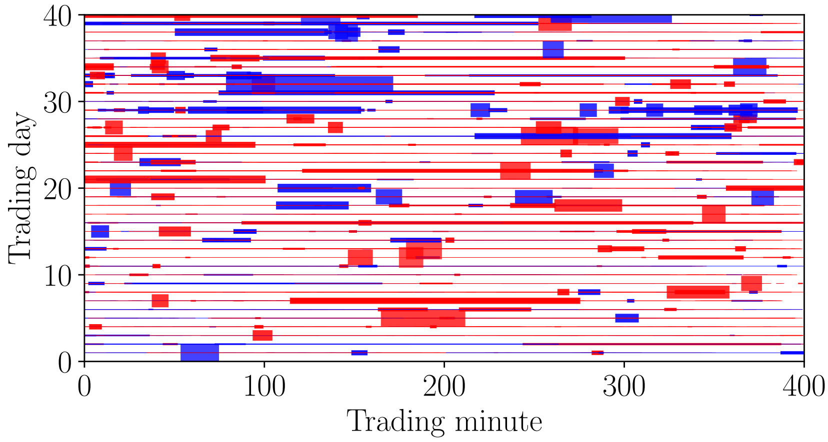

Figure 3.1: Time series of metaorders active on the U.S. equity market for MSFT in the period from June to July 2009. Buy (sell) metaorders are depicted in blue (red). The tickness of the line is proportional to the metaorder participation rate η. More metaorders in the same time interval implies darker colors. Each horizontal lines represents a fixed trading day and as evident there is almost always an active metaorder from our database, which is of course only a subset of the number of orders executed in the market.

traded in whole day by the market. It is easily seen that independently of the total daily volume, the volume time defined in that way equals to 0 at the market opening and to 1 at the market close.

In this metric the statistical properties of a metaorder of Q shares executed over a time interval $[t_s, t_e]$ can be described by the following three observables: the participation rate η, the volume duration D, and the unsigned order size ϕ. The participation rate η is defined as the ratio between the number of shares Q traded by the metaorder and the whole market volume during the execution interval $[t_s, t_e]$

$$\eta = \frac{Q}{V(t_e) - V(t_s)}. \tag{3.1}$$

The duration D expressed in volume time is equal to

$$D = \frac{V(t_e) - V(t_s)}{V_d}, \tag{3.2}$$

while the unsigned daily fraction is defined as the ratio between the metaorder unsigned volume Q and the volume V_d traded by the market in the whole day, i.e.

$$\phi = \frac{Q}{V_d} = \eta \times D. \tag{3.3}$$

Since all the metaorders are limited to a single trading day, by construction all these parameters are limited between 0 and 1.

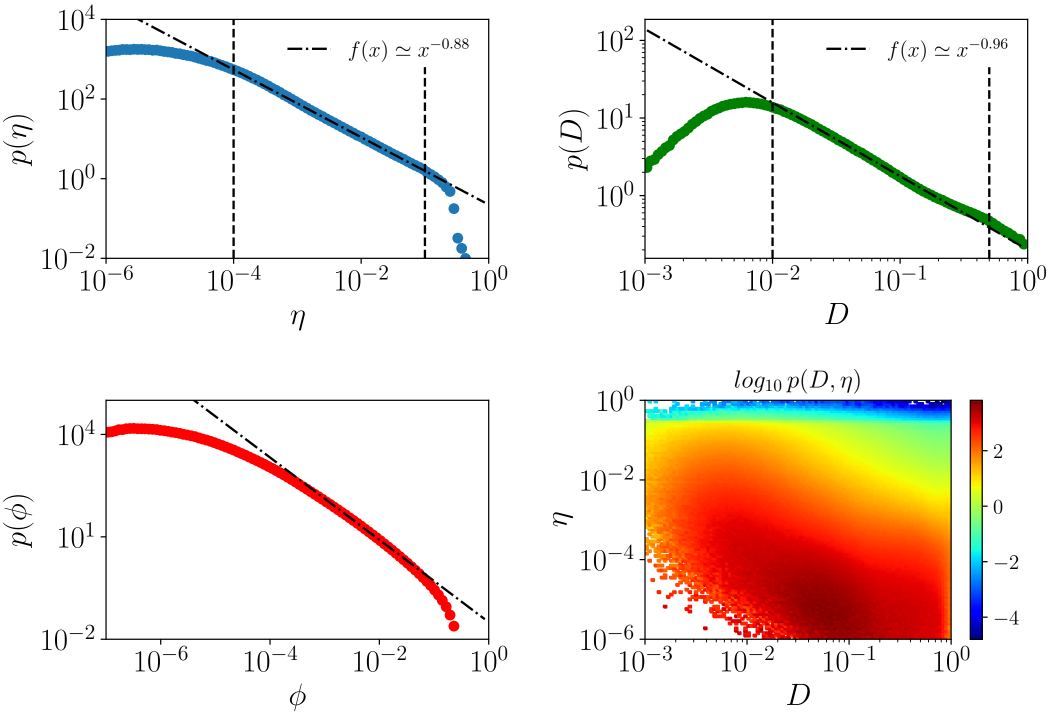

Figure 3.2: Estimation of the probability density function for the participation rate η (top left), volume duration D (top right), and unsigned daily fraction ϕ (bottom left). Note that all the panels are in log-log scale. The top panels also show the best fit with a power-law function in the region bounded by the vertical dashed lines. The bottom right panel shows the logarithm of the estimated joint probability density function $p(D, \eta)$ in double logarithmic scale of the duration D and of the participation rate η.

The salient statistical properties of respectively the participation rate η, the duration D and the order size ϕ, are illustrated in Figure 3.2. Firstly we find that the participation rate η and the duration D are both well approximated by truncated power-law distributions over several orders of magnitude. The estimated probability density function of the participation rate η is shown in log-log scale in the top left panel of Figure 3.2. A power-law fit in the region $10^{-4} \leq \eta \leq 10^{-1}$, i.e. over three orders of magnitude gives a best fit exponent $a = -0.882 \pm 0.001$. The top right panel of Figure 3.2 shows the estimated probability density function of the duration D of a metaorder. A power-law fit in the region bounded by vertical dashed lines ($0.01 \leq D \leq 0.5$) gives a power-law exponent $a = -0.964 \pm 0.002$. These power laws are very heavy-tailed, meaning that there is substantial variability in both the participation rate and the duration. Note that in both cases the variability is intrinsically bounded, and therefore the power law is automatically truncated: in fact by definition $\eta \leq 1$ and $D \leq 1$. In addition, for $p(D)$ there is a small bump on the right extreme of the distribution corresponding to metaorders executed during the whole trading day. Note that the deviation from a power law for small D is a consequence of our Filter 3 retaining only metaorders lasting at least 2 min, which in volume time corresponds in average to $2/390 \simeq 0.005$. The bottom left panel shows the probability density function of the unsigned daily fraction ϕ. In this case the distribution is less fat-tailed, and clearly not power-law. This is potentially an important result, as the predictions of some theories for market impact depend on this, and have generally assumed power-law behavior [14, 15]. Furthermore, the distribution of the order size ϕ is invariant respect to the metaorder direction (not shown). Finally, the bottom right panel of Figure 3.2 shows the logarithm of the estimated joint probability density function $p(D, \eta)$ in double logarithmic scale. We measure a very low linear correlation (-0.08) between the two variables, the main contribution coming from the extreme regions, i.e. very large η implies very small D and vice versa. In other words, as expected, very aggressive metaorders are typically short and long metaorders more often have a small participation rate. All these results are in agreement with what found by Zarinelli et al. in [10].

3.3 Possible measurements bias

In general it is rare to have access to a rich and detailed dataset as the one provided by ANcerno. It is then necessary to work with incomplete and imprecise data which can implies artifacts and biases. In such a way to obtain reproducible and understandable results, it is common practice to take in mind the following possible erroneous effects:

1. *Conditioning bias*: Larger the size Q of a metaorder, more probably it is generated from a stronger prediction signal. Therefore, in this case it is possible that market impact may not reveal any structural correlation with the order flow while may be due to short-term predictability.

2. *Prediction bias*: Traders who follow strong short-term price prediction signals may choose to execute their metaorders particularly quickly, to make more profit from their signal. Therefore, the strength of a prediction signal may itself influence the subsequent impact path, in particular when the prediction horizon is comparable to the execution horizon.

3. *Implementation bias*: In general it is reasonable to suppose that both the volume Q and the execution horizon T are fixed before a metaorder's execution starts. However, it is possible in reality that some asset manager may adjust these values during the execution over multiples days. For example, it can happen that a buy metaorder is only executed if the price goes down, and abandoned if the price goes up.

4. *Issuer bias*: Another possible bias may occur if a trader submits several consecutive and dependent metaorders successively. If such metaorders are positively correlated and occur close to one another in a time window, the impact of the first metaorder will be different to the impact of the subsequent metaorders.

5. *Synchronization bias*: The impact of a metaorder can change accordingly to whether or not other traders are seeking to execute similar metaorders at the same time. This can occur if different traders prediction signals are similar and correlated, or if they trade based on the same piece of information.

In general, most of them can be avoided when one has access to proprietary trading data containing detailed information on the execution metaorder. However, also for even less-ideal dataset these biases can statistically average out if the sample is sufficiently large.

Chapter 4

Crossover from linear to square root market impact

4.1 Introduction

In the literature, there are two different ways to modeling financial markets which greatly differ in their underlying assumptions. The first method, mostly advocated by financial mathematicians and economists, consists in postulating several global principles such as the martingale condition for market prices[1] and no arbitrage rules or perfect competition between market participants. One then seeks for a theory consistent with these axioms. While this method has a remarkable success it is also thouroughly insatisfactory since it fails to predict how economic agents behave such that price efficiency or no arbitrage is enforced. Hence, interest has shifted towards theories that put the emphasis on microscopic decision rules and focused on the consequences of the collective behaviour of the economic agents. In this second method, all events of the order book dynamics, represented by the setting and cancellation of limit orders and the triggering of market orders, occur according to a certain probabilistic rule, which may or may not depend on the past history. Due to their generality and simplicity the interest in these microscopic models has considerably increased in recent years.

In this research strand reaction-diffusion models have taken growing importance for the order book modeling starting from the highly stylized model proposed by Bak et al. [17, 74]. In a financial context it is supposed that two species of particles $\mathcal{A}$ and $\mathcal{B}$ represent the orders respectively on the ask and on the bid sides. These orders diffuse on the one dimensional grid of prices and when they meet they annihilate according to the rule $\mathcal{A}+\mathcal{B} \to \emptyset$ - a transaction at the mid-price occurs. The boundary between the $\mathcal{A}$-rich region and the $\mathcal{B}$-rich region corresponds to the mid-price S_t. Although in the preliminar reaction-diffusion models prices are sub-diffusive and market efficiency is not assured, recent research has shown that reaction-diffusion models combined with the idea of *latent order book* - an ideal and not measurable book where all the intentions of the

[1]The martingale condition implies that the conditional expectation of the future price, given the prior values, is equal to the present value.

traders are kept - allow to explain one of the misterious riddles in quantitative finance: the square-root law of metaorders [2, 7]. This idea opens the door to the development of a physics inspired *latent limit order book* (LLOB) model for the coarse-grained dynamics of latent liquidity [2, 7] which naturally explains why the impact of metaorders grows like the square-root of its size in a certain regime of parameters [7]. In a nutshell, the LLOB model predicts a linearly growing equilibrium liquidity profile, which implies a square-root impact law even in the absence of any reaction of the liquidity providers to the incoming metaorder. But this LLOB model also suggests that for a given execution time T, the very small Q regime should revert to a linear behaviour: the theory in fact predicts the detailed shape of the crossover between a linear to a square-root impact regime.

In the present era financial markets sputter enormous amounts of data that can be used to test scientific theories at levels of precision comparable to those achieved in physical sciences (see [91] for a recent example). In this light the aim of the present chapter is to test for the first time the detailed theoretical predictions of the crossover from a linear to a square-root impact using the large ANcerno database of metaorders, executed on the U.S. equity market and issued by a diversified set of institutional investors. We find that the crossover between linear and square-root impact is well described by the LLOB theory, albeit the transaction volume at the crossover point is much smaller than the predicted one. In fact, we argue that this can be accounted for by the coexistence of *slow* and *fast* agents in financial markets. *Fast* agents contribute to the total transaction volume but are unable to offer resistance against the execution of large metaorders. Therefore, only *slow* agents are able to dampen market impact and only their contribution is relevant for shaping up the square-root law. We recall how the LLOB model can be augmented to account for multiple agent frequencies, and compute the impact crossover function within this extended framework, resulting in a remarkably good fit of the data.

4.2 Dynamics of the latent limit order book

In this section we recall the LLOB model introduced in Donier et al. [7] since the theoretical part of our work leverages on it. This model provides a micro-structural explanation of the square-root law taking present its insensitivity to the high frequency dynamics which suggests that its interpretation should lie in some general properties of the low frequency, large scale dynamics of liquidity. The basic assumption underlying the arguments is the existence of a slowly evolving latent order book storing the volume that market participants would be willing to trade at any given price level [2]. In other words, this latent order book is where the true liquidity of the market lies, at variance with the real limit order book (LOB) where only a very small fraction of this liquidity - which evolves on very fast time scales - is revealed. This hypothesis is motivated by financial market data, which demonstrates that a tiny fraction of the daily volume traded on market is instantly available in the real order book. In fact, the publicly displayed liquidity at any given time is usually very small - typically on the order of 10^{-2} of the total daily transaction volume in stock markets. This is coherent with the fact that financial markets are the arena of a collective hide-and-seek game between buyers and sellers, resulting in a somewhat paradoxical situation where the total quantity that markets participants intend to trade is very large (0.5% of the total market capitalisation changes hands every day in stock markets) while most of this liquidity remains hidden, or latent. In fact, traders tend to hide their intentions as long as they can, as they have no incentive in giving away private information too soon by adding orders to the real order book: the actual decision to trade at a certain price in the future could be itself latent. This is the main reason why the vast majority of the daily traded volume progressively reveals itself as trading proceeds or in other words that liquidity is a dynamical process [34].

In the LLOB model it is assumed that each market participant trades randomly and independently intention on the latent order book in the same spirit of the zero-intelligence model used for the real order book by Smith et al. [75]. In this setup the fundamental quantities of interest are the average buy density $\varphi_{\mathcal{B}}(x,t)$ and the average sell density $\varphi_{\mathcal{A}}(x,t)$ of latent orders around the price x at time t. As argued in [2, 7], the coarse-grained dynamics of the latent liquidity close to the current mid-price S_t is described by the following system of coupled continuous reaction-diffusion equations:

$$\partial_t\varphi_{\mathcal{A}}(x,t) = -V_t\partial_x\varphi_{\mathcal{A}}(x,t) + \mathcal{D}\partial_{xx}\varphi_{\mathcal{A}}(x,t) - \nu\varphi_{\mathcal{A}}(x,t) + \lambda\Theta(S_t - x) - R_{\mathcal{A},\mathcal{B}}(x,t), \quad (4.1)$$

$$\partial_t\varphi_{\mathcal{B}}(x,t) = \underbrace{-V_t\partial_x\varphi_{\mathcal{B}}(x,t) + \mathcal{D}\partial_{xx}\varphi_{\mathcal{B}}(x,t)}_{\text{Drift-Diffusion}} - \underbrace{\nu\varphi_{\mathcal{B}}(x,t)}_{\text{Cancellation}} + \underbrace{\lambda\Theta(x - S_t)}_{\text{Deposition}} - \underbrace{R_{\mathcal{A},\mathcal{B}}(x,t)}_{\text{Reaction}}, \quad (4.2)$$

where the different terms in the right hand sides of Eqs. 4.2 and 4.1 represents respectively

- **Drift-Diffusion**: market participants can change their mind along time moving their intentions on the x-axis price. The drift term $V_t\partial_x$ represents a time dependent collective motion originated by an exogeneous source as for example a new piece of information. The diffusion term $\mathcal{D}\partial_{xx}$ describes instead the independent random motion of the trader latent orders that are thought as zero-intelligence agents [75].

- **Cancellation**: agents might remove completely or only partially trade intentions from the latent order book. This is modeled by the parameter ν which is assumed to be independent of the price level x and it corresponds to a memory time scale[2] equal to ν^{-1}.

- **Deposition**: the appearance of new buy/sell intentions is modeled by $\lambda\Theta(x)$ where λ is the rain intensity of new orders and $\Theta(x)$ is the Heaviside function, i.e. $\Theta(x > 0) = 1$ and $\Theta(x < 0) = 0$; note that unlike the cancellation term, the deposition intensity is independent on the density of the latent order book.

[2]In order of magnitude, it is a reasonable to suppose that the latent order book has a memory time scale from several hours to several days [2]. The original LLOB model considers only slow actors, i.e. institutional investors, neglecting the contribution of the high frequency traders.

- **Reaction**: when two orders meet at the same price x they annihilate according to the reaction $\mathcal{A}+\mathcal{B} \rightarrow \emptyset$ with an intensity proportional to a reaction rate k and to the product of the buy and sell densities, i.e. $R_{\mathcal{A},\mathcal{B}}(x,t) = k\varphi_{\mathcal{A}}(x,t)\varphi_{\mathcal{B}}(x,t)$. In the following we assume the limit $k \rightarrow \infty$ for which buy and sell orders annihilate instantaneously when they are at the same price. This corresponds to a latent order book without any overlap and with a transaction price S_t well defined by the condition
$$\varphi_{\mathcal{A}}(S_t,t) - \varphi_{\mathcal{B}}(S_t,t) = 0. \tag{4.3}$$

Although the dynamics of the latent densities $\varphi_{\mathcal{A}}(x,t)$ and $\varphi_{\mathcal{B}}(x,t)$ are not trivial as a consequence of the non-linearity introduced by the reaction term $R_{\mathcal{A},\mathcal{B}}(x,t)$, the combination
$$\varphi(x,t) = \varphi_{\mathcal{B}}(x,t) - \varphi_{\mathcal{A}}(x,t) \tag{4.4}$$
satisfies the following differential equation independent from the reaction term
$$\partial_t\varphi(x,t) = -V_t\partial_x\varphi(x,t) + \mathcal{D}\partial_{xx}\varphi(x,t) - \nu\varphi(x,t) + \lambda \ \mathrm{sign}(S_t - x) \tag{4.5}$$
and it allows to recover the bid and the ask densities respectively as follows
$$\varphi_{\mathcal{A}}(x,t) = -\varphi(x,t)\Theta(x - S_t), \tag{4.6}$$
$$\varphi_{\mathcal{B}}(x,t) = \varphi(x,t)\Theta(S_t - x). \tag{4.7}$$
Removing the drift term in Eq. (4.5) through the change of variable $y = x - \hat{S}_t$ with $\hat{S}_t = \int_0^t V_\tau \, d\tau$ the dynamics of the latent density $\varphi(x,t)$ is described by
$$\partial_t\varphi(y,t) = \mathcal{D}\partial_{yy}\varphi(y,t) - \nu\varphi(y,t) + \lambda\,\mathrm{sign}(S_t - \hat{S}_t - y). \tag{4.8}$$
Note that the change of variable $y = x - \hat{S}_t$ is a mathematical trick which allows to work in the reference frame of the fundamental price $\hat{S}_t$ where only the endogeneous price changes are relevants: in other words, the deterministic evolution of the LLOB is treated independently from the random dynamics of the price.

From Eq. (4.8) we can then investigate the price dynamics $y_t = S_t - \hat{S}_t$ during the execution of a metaorder imposing the following boundary conditions
$$\varphi(y_t,t) = 0 \quad \forall t, \tag{4.9}$$
$$\lim_{|y|\rightarrow\infty}\varphi(y,t) \neq \infty, \tag{4.10}$$
$$\varphi(y,t=0) = -\varphi(-y,t=0), \tag{4.11}$$
with $\varphi(y,t=0)$ the initial symmetric state of the latent order book.

4.3 Insight on the stationary state

In an initial symmetric latent order book, i.e. $\varphi(y, t = 0) = -\varphi(-y, t = 0)$ and $S_{t=0} = \hat{S}_{t=0}$, by symmetry it follows that $S_t = \hat{S}_t$ for all times $t > 0$. The stationary solution $\varphi^{\mathrm{st}}(y)$ is then given by setting $\partial_t \varphi^{\mathrm{st}}(y, t) = 0$ in Eq. (4.8)

$$\mathcal{D}\partial_{yy}\varphi^{\mathrm{st}}(y) - \nu\varphi^{\mathrm{st}}(y) = \lambda \tag{4.12}$$

which implies

$$\varphi^{\mathrm{st}}(y) = -\frac{\lambda}{\nu}\mathrm{sign}(y)(1 - e^{-\gamma|y|}) \tag{4.13}$$

with $\gamma := \sqrt{\nu/\mathcal{D}}$ the inverse of the typical length scale below which the latent order book is locally linear, i.e.

$$\varphi^{\mathrm{st}}(y) \approx -\frac{\lambda\gamma}{\nu}y = -\mathcal{L}y. \tag{4.14}$$

The coefficient $\mathcal{L} := \lambda/\sqrt{\nu\mathcal{D}}$ in Eq. (4.14) is interpretable as a measure of the latent liquidity in the market and it enters in the definition of the total transaction rate J (the flux of orders through the origin) given by

$$J := \mathcal{D}|\partial_y\varphi^{\mathrm{st}}(y)|_{y=0} = \mathcal{D}\mathcal{L}. \tag{4.15}$$

4.4 Square root impact within the latent linear order book

Let us discuss now how the square-root law comes out in the infinite memory limit, namely for $\nu, \lambda \to 0$ with $\mathcal{L} \sim \lambda\nu^{-1/2}$ constant, when a metaorder with trading intensity rate $m(t)$ and duration T is executed in an inital locally linear latent order book $\varphi(y, t = 0) = \varphi^{\mathrm{st}}(y)$ [3]. In this setup a metaorder with volume $Q = \int_0^T m(\tau)d\tau$ is modeled by an extra term at the mid-price, i.e.

$$\partial_t\varphi(y, t) = \mathcal{D}\partial_{yy}\varphi(y, t) + m(t)\delta(y - y_t). \tag{4.16}$$

The price dynamics is then derived solving Eq. (4.16) with the boundary condition

$$\lim_{y\to\infty} \partial_y\varphi(y, t) = -\mathcal{L} \tag{4.17}$$

which is equivalent to assume that far from the mid-price the latent order book refill at a constant rate. Moving to the Fourier space Eq. (4.16) can be rewritten as

$$\partial_t\tilde{\varphi}(k, t) = -\mathcal{D}k^2\tilde{\varphi}(k, t) + m(t)e^{iky_t} \tag{4.18}$$

where setting $\tilde{\varphi}(k, t) := \tilde{F}(k, t)e^{-\mathcal{D}k^2t}$ we obtain that

$$\tilde{F}(k, t) = \tilde{F}(k, 0) + \int^t d\tau\, m(\tau)e^{\mathcal{D}k^2\tau + iky_\tau}. \tag{4.19}$$

[3] It is assumed that the metaorder is small enough in such a way to not change the market model parameters $\mathcal{D}$, ν and λ.

Antitransforming it back gives

$$\varphi(y,t) = -\mathcal{L}y + \int_0^t d\tau \frac{m(\tau)}{\sqrt{4\pi\mathcal{D}(t-\tau)}} e^{-\frac{(y-y_\tau)^2}{4\mathcal{D}(t-\tau)}} \tag{4.20}$$

and imposing $\varphi(y_t, t) = 0$ we obtain the following self-consistent relation for the transaction price

$$y_t = \frac{1}{\mathcal{L}} \int_0^t d\tau \frac{m(\tau)}{\sqrt{4\pi\mathcal{D}(t-\tau)}} e^{-\frac{(y_t-y_\tau)^2}{4\mathcal{D}(t-\tau)}}. \tag{4.21}$$

Although, in general Eq. (4.21) can be solved only numerically some insights on its solution can be derived analytically in the case of a metaorder executed with a constant trading rate $m_0 = Q/T$. In fact, taking present that the argument of the exponential term in Eq. (4.21) can be rewritten as follows

$$\begin{aligned}\frac{(y_t - y_\tau)^2}{4\mathcal{D}(t-\tau)} = \frac{1}{4\mathcal{D}(t-\tau)} \Big(\frac{1}{\mathcal{L}} \int_0^t du \frac{m_0}{\sqrt{4\pi\mathcal{D}(t-u)}} e^{-\frac{(y_t-y_u)^2}{4\mathcal{D}(t-u)}} + \\ - \int_0^\tau d\tau \frac{m_0}{\sqrt{4\pi\mathcal{D}(\tau-u)}} e^{-\frac{(y_\tau-y_u)^2}{4\mathcal{D}(\tau-u)}} \Big)^2 \propto \Big(\frac{m_0}{\mathcal{D}\mathcal{L}}\Big)^2 = \Big(\frac{m_0}{J}\Big)^2\end{aligned} \tag{4.22}$$

it emerges that the behaviour of the market impact $I := y_T$ depends on the participation rate $\eta = m_0/J$. For this reason let us focus in Eq. (4.21) on the two limit regimes respectively with small ($\eta \ll 1$) and large ($\eta \gg 1$) participation rates before to discuss its numerical solution.

4.4.1 Small participation rate regime

In the limit of small participation rate $\eta \ll 1$ Eq. (4.22) justifies the use of the following approximation $e^{-\frac{(y_t-y_\tau)^2}{4\mathcal{D}(t-\tau)}} \approx 1$ in Eq. (4.21). This implies that

$$y_t = \frac{1}{\mathcal{L}} \int_0^t d\tau \frac{m_0}{\sqrt{4\pi\mathcal{D}(t-\tau)}} = \frac{m_0}{\mathcal{L}} \sqrt{\frac{t}{\pi\mathcal{D}}} = \sqrt{\frac{\mathcal{D} m_0 Q_t}{\pi J^2}} \tag{4.23}$$

which exactly boils down to the linear propagator model proposed in [18], i.e.

$$y_t = \int_0^t d\tau\, \mathcal{G}(t-\tau)\, m_0 \tag{4.24}$$

with the propagator equal to $\mathcal{G}(t-\tau) \sim (t-\tau)^{-1/2}$.

4.4.2 Large participation rate regime

In the limit of large participation rate $\eta \gg 1$ we can solve Eq. (4.21) through the combination of the variable change $u = t - \tau$ and of the saddle point approximation. In fact, in the limit of $u \to 0$ the exponential term in Eq. (4.21) goes to zero faster than everything and since $y_t = y_{t-u} + \dot{y}_t\, u + o(u)$ it follows that

$$y_t = \frac{1}{\mathcal{L}} \int_0^t du \frac{m_0}{\sqrt{4\pi \mathcal{D} u}} e^{-\frac{(y_t - y_{t-u})^2}{4\mathcal{D}u}} \approx \frac{1}{\mathcal{L}} \int_0^\infty du \frac{m_0}{\sqrt{4\pi \mathcal{D} u}} e^{-\frac{\dot{y}_t^2 u}{4\mathcal{D}}} = \frac{m_0}{\mathcal{L}\dot{y}_t \sqrt{\pi}} \int_0^\infty dv \frac{e^{-v}}{\sqrt{v}} = \frac{m_0}{\mathcal{L}\dot{y}_t}. \tag{4.25}$$

Note that in the second passage we extended the integral up to infinity, i.e. $\int_0^t du \cdots \to \int_0^\infty du \cdots$ - which is a good approximation taking present that the exponential term is around zero - in such a way to recover the definition of the Γ function. Since y_t must be an increasing function on time it follows from Eq. (4.25) that

$$\frac{1}{2}\partial_t y_t^2 = y_t \dot{y}_t = \frac{m_0}{\mathcal{L}} \tag{4.26}$$

which implies that

$$y_t = \sqrt{\frac{2\mathcal{D} m_0 t}{J}}. \tag{4.27}$$

4.4.3 Numerical solution

In general Eq. (4.21) can be solved only numerically. In the case of a constant trading intensity rate $m_0 = Q/T$ it follows that the price impact $I := y_T$ is described by

$$I(Q) = \sqrt{\frac{\mathcal{D}Q}{J}}\, \mathcal{F}(\eta)\ , \tag{4.28}$$

with $\eta = Q/(JT)$ the participation rate and $\mathcal{F}(\eta)$ a scaling function (see Figure 4.1) represented by

$$\mathcal{F}(\eta) \approx \begin{cases} \sqrt{\eta/\pi} & \text{for } \eta \ll 1 \quad \text{small participation rate regime} \\ \sqrt{2} & \text{for } \eta \gg 1 \quad \text{large participation rate regime.} \end{cases} \tag{4.29}$$

It emerges that in the LLOB model the market impact $I(Q)$ is predicted to be linear in Q for small Q at fixed T, and crosses over to a square root for large Q. Furthermore, in the square root regime the market impact is predicted to be *independent* from the execution time T.

4.5 Empirical analysis

We now turn to the ANcerno database to see how well Eq. (4.28) is supported empirically. Our sample covers a total of 880 trading days, from January 2007 to June 2010, and we follow the cleaning procedure introduced in [10] to remove possible spurious effects. The sample is represented by around 8 million metaorders uniformly distributed in time and market capitalization[4]. Each metaorder in the database is characterised by a broker

[4]The sample represents around the 5% of the total market volume

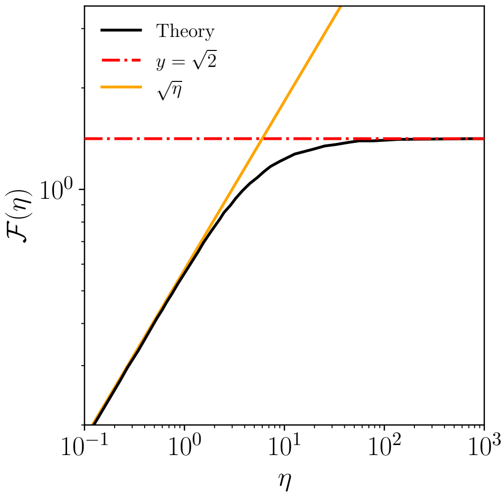

Figure 4.1: (Black solid line) Dependence on the participation rate η of the scaling function $\mathcal{F}(\eta)$ (see Eq. (4.28)) computed in the case of a constant trading intensity $m_0 = Q/T$. The curve interpolates between a $\sqrt{\eta}$ dependence observed at small participation rate η (orange solid line) and an asymptotlically constant regime $\simeq \sqrt{2}$ for large participation rate η (red dashed line).

label, a stock symbol, the total number of traded shares Q, the sign $\epsilon = \pm 1$ (buy/sell), the start-time t_{s} and the end-time t_{e} of its execution. In line with the definition given above, and following [10], the participation rate is measured as $\eta = Q/V_T$ where $V_T = V(t_{\mathrm{e}}) - V(t_{\mathrm{s}})$ is the total volume traded in the market during the metaorder execution. In order to compare different stocks with very different daily volumes, we measure Q in units of the corresponding daily volume V_{d}, and introduce the volume fraction $\phi := Q/V_{\mathrm{d}}$, which in the model notation is equal to Q/JT_{d}, with $T_{\mathrm{d}} = 1$ day. We will also measure execution in relative volume time and redefine the execution time as $D := (V(t_{\mathrm{e}}) - V(t_{\mathrm{s}}))/V_{\mathrm{d}}$. Finally, we introduce rescaled log-prices as $s_t := (\log S_t)/\sigma_{\mathrm{d}}$, where $\sigma_{\mathrm{d}} = (S_{\mathrm{high}} - S_{\mathrm{low}})/S_{\mathrm{open}}$ is the daily volatility estimated from the daily high, low and open prices.

The average price impact $I(Q)$ for a given executed volume Q, as studied in most previous studies [9, 10, 35, 89], is defined as:

$$I(Q) = \mathbb{E}[\epsilon \cdot (s_{\mathrm{end}} - s_{\mathrm{start}})|Q] \tag{4.30}$$

where $s_{\mathrm{start}} = s_{t_{\mathrm{s}}}$, $s_{\mathrm{end}} = s_{t_{\mathrm{e}}}$ are, respectively, the rescaled logarithmic mid-price at the start and at the end of the metaorder.

In order to test directly Eq. (4.28), we estimate the scaling function $\mathcal{F}(\eta)$ by dividing the data into evenly populated bins of constant participation rate η and computing the conditional expectation of $\epsilon(s_{\mathrm{end}} - s_{\mathrm{start}})/\sqrt{\phi}$ for each bin in participation rate.

According to the LLOB model this expectation is equal to $\sqrt{\mathcal{D}/\sigma_{\rm d}^2}\,\mathcal{F}(\eta)$. Here and in the following, error bars are determined as standard errors.

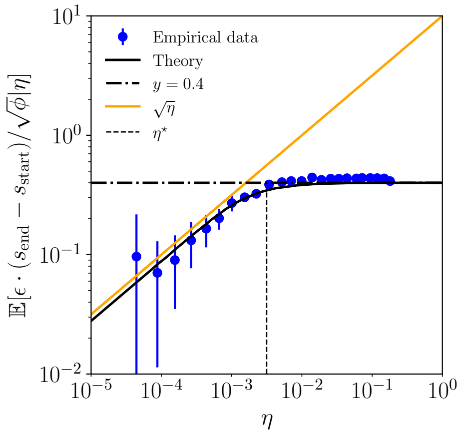

Figure 4.2: Empirically determined scaling function $\mathcal{F}(\eta)$ vs. participation rate η. The data (blue points) interpolates between a $\sqrt{\eta}$ behaviour observed at small participation rates and an asymptotically constant regime ≈ 0.4 for large η, i.e. for $\eta \gtrsim \eta^\star$ with $\eta^\star \approx 3.15 \times 10^{-3}$. Black line: prediction of the LLOB model, with an adjusted crossover $\eta^\star := J_{\rm s}/J_{\rm f}$ allowing for the existence of two categories of agents (*fast* and *slow*). The data points are obtained by restricting to metaorders with sufficient large order size, i.e. $\phi \gtrsim 10^{-5}$.

The results are shown in Figure 4.2 and are, up to a rescaling of both the axis, remarkably well accounted for by the LLOB function $\mathcal{F}(\eta)$ (illustrated in Figure 4.1), that describes the crossover between a linear-in-Q regime for small participation rates η, and a T-independent, $\sqrt{Q}$ regime at large η. Whereas a linear regime for small Q's was already reported in [10, 41], the scaling analysis provided here has not been attempted before. In fact, deviations from a pure square-root were observed in [10] where the authors fitted the data with a logarithmic function $\ln(a + b\,Q)$, which indeed behaves linearly for small arguments. The fact that impact in $\sqrt{Q}$ regime chiefly depends on Q but not on T is compatible with the results of [10], but contradicts theories that assigns the $\sqrt{Q}$ dependence to *duration* of the metaorder, as in [14, 36, 95]. In Appendix A we discuss in details this independence to the duration T arguing that the market impact should not be confused with price diffusion, i.e. that typical price changes grow as the square root of T. Furthermore, directly regressing the impact as $\sqrt{\phi}T^{-\beta}$ in the $\eta > \eta^\star$ regime yields $\beta = -0.04 \pm 0.02$, confirming the near independence of the market impact on the execution time T; while in the $\eta < \eta^\star$ regime the regression as $\phi T^{-\beta}$ yields $\beta = 0.45 \pm 0.05$, in close agreement with the LLOB prediction $\beta = 1/2$.

However, whereas the crossover between the two regimes should occur around $\eta^\star = 1$ within the original LLOB model (see Figure 4.1), empirical data points towards a much smaller value $\eta^\star \sim 10^{-3}$. This is actually consistent with the fact that all the empirical evidence for the square-root law reported in the literature concern moderate participation rates (typically in the range $10^{-3} - 10^{-1}$, see e.g. [4, 12, 38, 41]) but never in a regime where the volume of the metaorder becomes larger than that of the rest of the market, as would be requested within the LLOB specification. Note that in our sample, 70% of the metaorders are such that $\eta > \eta^\star$.

4.6 Fast and slow latent order books

In order to account for this large discrepancy in the value of $\eta^\star$, we shall consider the extended LLOB model recently proposed by Benzaquen et al. [69] to include agents with different time horizons, as it is clearly the case in financial markets. In the simplest case of a bi-modal distribution of agents (*fast* and *slow*), the LLOB formalism can be generalized to describe two latent order book densities, respectively $\varphi_{\mathrm{s}}(y,t)$ for the *slow* liquidity and $\varphi_{\mathrm{f}}(y,t)$ for the *fast* liquidity – for example provided by high frequency traders. The corresponding dynamical equations then read [69]:

$$\partial_t \varphi_\circ(y,t) = \mathcal{D}_\circ \partial_{yy} \varphi_\circ(y,t) - \nu_\circ \varphi_\circ(y,t) + \lambda_\circ \mathrm{sign}(y_t) + m_\circ(t)\delta(y - y_t) \ , \tag{4.31}$$

where $\circ = \mathrm{s}$, f and $m_{\mathrm{s}}(t)$ (resp. $m_{\mathrm{f}}(t)$) is the fraction of the metaorder absorbed by the *slow* (resp. *fast*) traders, with $m_{\mathrm{s}}(t) + m_{\mathrm{f}}(t) = m_0$. To note that we allow the activity rate of the two categories of agents to be different through the coefficients $\mathcal{D}_\circ$, $\nu_\circ$ and $\lambda_\circ$. The interesting limits for our purpose are:

- $J_{\mathrm{s}} \ll J_{\mathrm{f}}$ and $m_0 \ll J_{\mathrm{f}}$, where $J_\circ = \lambda_\circ \sqrt{\mathcal{D}_\circ/\nu_\circ}$. These inequalities mean that (i) the total transaction rate $J = J_{\mathrm{s}} + J_{\mathrm{f}} \approx J_{\mathrm{f}}$ is dominated by *fast* traders and (ii) the flux corresponding to the metaorder is small compared to the total transaction rate of the market, as with most metaorders executed in liquid markets.

- $\nu_{\mathrm{s}} T \ll 1$ and $\nu_{\mathrm{f}} T \gg 1$. As shown in [69] this implies that *slow*, persistent agents are able to resist to the impact of the metaorder, whereas *fast* agents are playing the role of *transparent* intermediaries, only lubricating the high-frequency activity of markets.

This double-frequency model can be solved exactly in some limits. One should distinguish two cases, depending on whether the execution time T is larger or smaller than a certain $T^\dagger := \nu_{\mathrm{f}}^{-1} \eta^{\star -2} \mathcal{D}_{\mathrm{s}}/\mathcal{D}_{\mathrm{f}}$, where $\eta^\star := J_{\mathrm{s}}/J_{\mathrm{f}}$. For $T > T^\dagger$, the scaling result Eq. (4.28) is simply modified as:

$$I(Q) = \sqrt{\frac{\mathcal{D}_{\mathrm{s}} Q}{J_{\mathrm{s}}}} \, \mathcal{F}\left(\frac{\eta}{\eta^\star}\right) . \tag{4.32}$$

For $T < T^\dagger$, this result is further multiplied by $\sqrt{T/T^\dagger}$, with a shifted crossover point $\eta^\star \to \eta^\star T^\dagger / T$. If we assume that $T^\dagger$ is small enough for all data points (which needs to be checked a posteriori), then the prediction of the double-frequency model, Eq. (4.32), is precisely the same as the one of the standard LLOB model, Eq. (4.28), up to a rescaling of the x-axis by $\eta^\star$, and of the y-axis by a ratio $\sqrt{\mathcal{D}_{\mathrm{s}}J/\mathcal{D}J_{\mathrm{s}}}$. Figure 4.2 shows that the LLOB scaling prediction indeed reproduces the data very well, which allows a direct determination of $\eta^\star \approx J_{\mathrm{s}}/J \approx 3.15 \times 10^{-3}$. In other words, we find that most of the daily liquidity is provided by *fast* agents, whereas the resistance to a metaorder relies on a small fraction of *slow* agents. We have checked that neither the quality of the fit nor the value of $\eta^\star$ are significantly different in the period 2007-2008 and 2009-2010 as evident in Figure 4.3. We have also investigated the dependence of $\eta^\star$ on market capitalization and volatility (see panels in Figure 4.4). We find that low volatility/large capitalization stocks are characterized by a larger value of $\eta^\star$ than high volatility/mid-small capitalization stocks, suggesting, perhaps counter-intuitively, that the low frequency activity is comparatively more important in low volatility/large capitalization stocks.

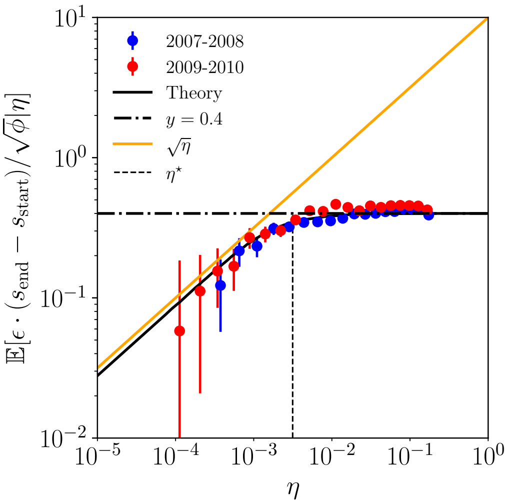

Figure 4.3: Empirically determined scaling function $\mathcal{F}(\eta)$ vs. participation rate η for metaorders in the period 2007-2008 (blue dots) and in the period 2009-2010 (red dots). The data (blue and red points) interpolates between a $\sqrt{\eta}$ behaviour observed at small participation rates and an asymptotically constant regime ≈ 0.4 for large η, i.e. for $\eta \gtrsim \eta^\star$ with $\eta^\star \approx 3.15 \times 10^{-3}$ independent of the selected time period. Black line: prediction of the *fast* and *slow* latent order books model as discussed in the main text. The data points are obtained by restricting to metaorders with sufficient large order size, i.e. $\phi \gtrsim 10^{-5}$.

The plateau value for $\eta > \eta^\star$, on the other hand, leads to $\sqrt{\mathcal{D}_{\mathrm{s}}J/\mathcal{D}J_{\mathrm{s}}} = 0.4/\sqrt{2}$, leading to $\sqrt{\mathcal{D}_{\mathrm{s}}/\mathcal{D}} \simeq 10^{-2}$. Since $\sqrt{\mathcal{D}}$ should be close to the price volatility [7], we find that, consistently with its interpretation, the *slow* liquidity moves much more slowly than the price itself. These estimates in turn lead to $T^\dagger \approx 45\,\nu_{\mathrm{f}}^{-1}$, or ~ 45 seconds for $\nu_{\mathrm{f}}^{-1} = 1$ second. Since the median execution time of the metaorders in our sample is 35

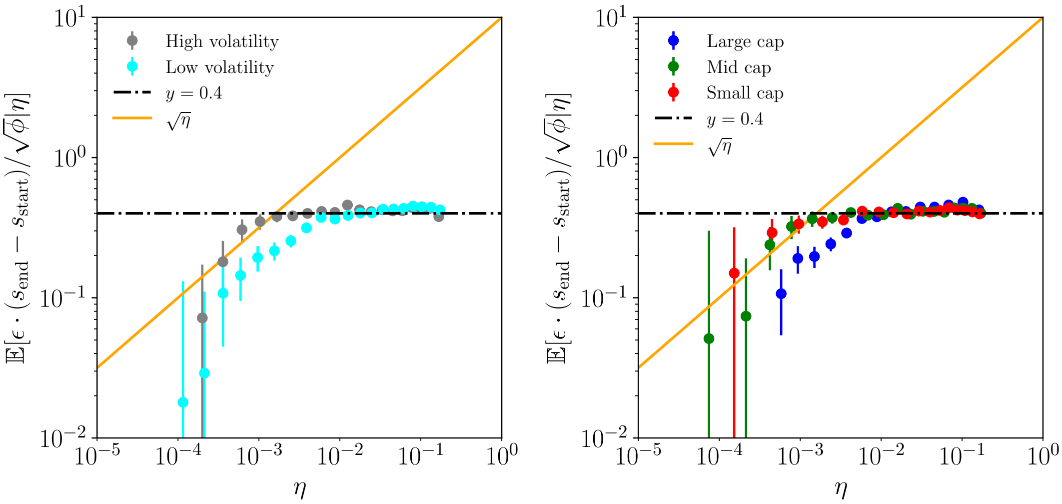

Figure 4.4: (Left panel) Empirically determined scaling function $\mathcal{F}(\eta)$ vs. participation rate η for metaorders executed on stocks with high volatility (grey dots) and small volatility (cyan dots). (Right panel) Empirically determined scaling function $\mathcal{F}(\eta)$ vs. participation rate η for metaorders with large (blue dots), mid (green dots), and small (red dots) capitalizations. In both the panels the data points are obtained by restricting to metaorders with sufficient large order size, i.e. $\phi \gtrsim 10^{-5}$. Note that low volatility/large capitalization stocks are characterized by a larger value of $\eta^\star$ than high volatility/mid-small capitalization stocks.

minutes, we conclude that most metaorders in our sample are indeed longer than $T^\dagger$.

Still, a bi-modal distribution of trading frequencies is certainly an oversimplification. One should consider instead, as in [69], a continuous distribution of frequencies (see Appendix 4.A.1 for a brief discussion on this point). Several empirical facts about the dynamics of financial markets (see e.g. [27, 34, 79]) actually suggest that such a distribution is a power-law. The numerical solution and the fitting procedur of such a general model is beyond the scope of the present chapter, but the simplified analysis in [69] suggests that the LLOB scaling function should be approximately valid, with a crossover value $\eta^\star$ that decreases as a power-law of D. Intuitively, the critical participation rate $\eta^\star$ should be larger for small metaorder duration D, since there are less traders that can be considered *fast* on such short time scales and more traders that are *slower* than D. This intuition is indeed confirmed by the main panel of Figure 4.5 where we show the rescaled data as a function of η, for metaoders longer and shorter than the median execution time $\bar{D} \approx 0.09$. The crossover participation rate $\eta^\star$ for small durations is found to be 10 times larger for large durations. In the inset of Figure 4.5, we show the D-dependence of $\eta^\star$, obtained by fitting the rescaled data by $\mathcal{F}(\eta/\eta^\star)$ using five bins of D containing the same number of data points ($\sim 1.4 \times 10^6$), suggesting $\eta^\star \sim 1/\sqrt{D}$.

It would be very interesting to use this result to map out the frequency distribution of the hidden liquidity, but this requires going beyond the approximate solution of [69]. We leave this for a future investigation. Another important feature that should be properly accounted for in a more detailed analysis is the *co-impact* effect discussed in Chapter 5, due to different metaorders executed during the same time period. But in a first approximation the sign correlation between metaorders simply lead in the LLOB framework to an effective increase of the order size ϕ.

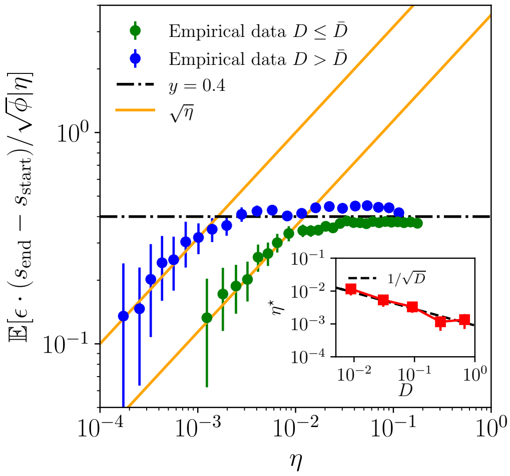

Figure 4.5: (Main panel) Empirically determined scaling function $\mathcal{F}(\eta)$ vs. participation rate η for metaorders with duration D larger (blue dots) or smaller (green dots) than the median sample duration $\bar{D} \approx 0.09$. Note that the crossover value $\eta^\star$ is ~ 10 times larger in the latter case. Note that the asymptotic value of $\mathcal{F}(\eta)$ (and hence the impact for a given Q) is approximately independent of D, as predicted by the LLOB theory. Both empirical curves are obtained using metaorders with sufficient large order size, i.e. $\phi \gtrsim 10^{-5}$. (Inset) Plot of the crossover participation rate $\eta^\star$ as a function of the execution duration D, revealing an approximate $1/\sqrt{D}$ behaviour.

4.7 Conclusions

In this chapter, we have used a very large data set of orders executed in the U.S. equity market to quantitatively test for the first time a recently proposed dynamical theory of liquidity that makes specific predictions about the shape of the market impact: a crossover from a linear (in volume) behaviour for small volumes to a square-root behaviour for intermediate volumes is predicted. The data unambiguously suggests the existence of such a crossover, and once again confirms the square-root law which, as emphasized on several previous occasions, is remarkably independent of the execution time – contradicting many early theories [14, 36, 95]. We have shown how the data points towards the existence of multiple time scales in the dynamics of liquidity, with its high frequency component dominating the total market activity and its low frequency component contributing to the concavity of the impact function: for a complementary viewpoint, see [40]. Our results are interesting from two rather different points of views. One is that they represent a significant improvement in our understanding of the determinants of market impact which is both the main component of trading costs for institutional investors and an important aspect of the stability of financial markets. The second aspect is that we are entering an era where economic and financial data becomes of such quality that theoretical ideas can be tested with standards comparable to those of natural sciences.

Appendix

4.A Beyond the infinite memory limit order book

In the original LLOB model [7] the authors focused on the infinite memory limit, namely ν, $\lambda \to 0$ while keeping $\mathcal{L} \sim \lambda\nu^{-1/2}$ constant, such that the latent order book becomes exactly linear in the proximity of the mid-price. This particular limit allows a simpler mathematical analysis of the market impact associated with the execution of a metaorder. However, since financial markets are characterized by agents with a broad spectrum of timescales, from low frequency institutional investors to high frequency traders, it is opportune to take in consideration the effects of non-vanishing cancellation ν and deposition λ rates for the dynamic of the latent order book [69]. This means that the latent order book dynamic is described by

$$\partial_t \varphi(y,t) = \mathcal{D}\partial_{yy}\varphi(y,t) - \nu\varphi(y,t) + \lambda\,\mathrm{sign}(S_t - y) \tag{4.33}$$

whose solution can be computed as

$$\varphi(y,t) = (\mathcal{G}_\nu \star \varphi^{\mathrm{st}})(y,t) + \int_{-\infty}^{\infty} d\xi \int_0^{\infty} d\tau\, \mathcal{G}_\nu(y-\xi, \xi-\tau)\, h(\xi,\tau) \tag{4.34}$$

where $h(\xi,\tau) = \lambda\,\mathrm{sign}(S_t - \xi)$ is the source term, $\varphi^{\mathrm{st}}(y)$ is the initial stationary state of the latent order book and $\mathcal{G}_\nu(y,t) = e^{-\nu t}\mathcal{G}(y,t)$ is the diffusion kernel with

$$\mathcal{G}(y,t) = \Theta(t)\frac{e^{-\frac{y^2}{4\mathcal{D}t}}}{\sqrt{4\pi\mathcal{D}t}} \tag{4.35}$$

and $\Theta(t)$ the Heaviside function. Assuming that a metaorder with intensity rate $m(t)$ and duration T is executed, the source term in Eq. (4.34) becomes $h(y,t) = m(t)\delta(y - S_t) + \lambda\,\mathrm{sign}(S_t - y)$. Performing the first integral in Eq. (4.34) we obtain

$$\varphi(y,t) = \varphi^{\mathrm{st}}(y)e^{-\nu t} + \int_0^{\min(t,T)} d\tau\, m(\tau)\, \mathcal{G}_\nu(y - S_\tau, t-\tau) - \lambda \int_0^t d\tau\, \mathrm{erf}\left[\frac{y - S_\tau}{\sqrt{4\mathcal{D}(t-\tau)}}\right] e^{-\nu(t-\tau)}. \tag{4.36}$$

Unfortunately Eq. (4.36) is not analytically tractable in the general case, but different interesting limit cases can be investigated assuming a constant trading intensity rate $m_0 = Q/T$. In particular, one may consider the following cases:

1. Small participation rate ($m_0 \ll J$) vs. large participation rate ($m_0 \gg J$).
2. Fast execution ($\nu T \ll 1$) vs. slow execution ($\nu T \gg 1$).
3. Small metaorder volumes $Q \ll Q_{\text{lin}}$ (for which the linear approximation of the stationary book is appropriate) vs. large volumes $Q \gg Q_{\text{lin}}$ (for which the linear approximation is no longer valid) with $Q_{\text{lin}} = J\nu^{-1}$.

	α	z_t^0	z_t^1
$m_0 \ll J$	$\frac{m_0}{\mathcal{L}\sqrt{\pi\mathcal{D}}}$	$\sqrt{t}$	$\left(\frac{\sqrt{\pi}}{2} - \frac{2}{\sqrt{\pi}}\right) t$
$m_0 \gg J$	$\sqrt{\frac{2m_0}{\mathcal{L}}}$	$\sqrt{t}$	$-\frac{1}{3}\sqrt{\frac{J}{2m_0}}\, t$

Table 4.1: Price tajectories for small ($m_0 \ll J$) and large ($m_0 \gg J$) participation rate regimes (see Eq. (4.37)).

In these limit cases the solution of the mid-price can be derived from Eq. (4.36) expanding y_t up to first order in $\sqrt{\nu}$ as

$$y_t = \alpha[z_t^0 + \sqrt{\nu} z_t^1 + o(\nu)], \tag{4.37}$$

where z_t^0 and z_t^1 denote respectively the zero and the first order contributions. Table 4.1 gathers the results for fast execution ($\nu T \ll 1$) and small metaorder volumes ($Q \ll Q_{\text{lin}}$) while the price trajectory for slow execution ($\nu T \ll 1$) and/or large metaorder volumes ($Q \gg Q_{\text{lin}}$) simply reads

$$y_t = \frac{m_0 \nu}{\lambda} t. \tag{4.38}$$

For a detailed derivation of these results we refer to the calculations discussed in [69].

4.A.1 Multi-timescales latent order book

The double-frequency framework previously discussed represents the starting point for a more realistic modelisation based on a continuous range of cancellation and deposition rates. For this extension it is necessary to solve an infinite system of equations indexed by the cancellation rate ν, i.e.

$$\partial_t \varphi_\nu(y,t) = \mathcal{D}_\nu \partial_{yy} \varphi_\nu(y,t) - \nu\varphi_\nu(y,t) + \lambda_\nu \text{sign}(y_\nu(t) - y) + m_\nu(t)\delta(y - y_\nu(t)) \tag{4.39}$$

where $\varphi_\nu(y,t)$ describes the contribution to the latent order book of agents with a typical cancellation frequency ν and deposition rate $\lambda_\nu = \mathcal{L}_\nu\sqrt{\nu\mathcal{D}_\nu}$. This system of equations must be then solved imposing the following boundary conditions

$$\int_0^\infty d\nu \rho(\nu) m_\nu(t) = m(t), \tag{4.40}$$

$$y_\nu(t) = y_t, \tag{4.41}$$

with $\rho(\nu)$ the distribution of cancellation rates and $m(t)$ the trading intensity rate of the metaorder. The interested reader can find in [69] the case based on the assumption of a power-law distribution of frequency.

Chapter 5

Co-impact: crowding effects in institutional trading activity

5.1 Introduction

The market impact of trades, i.e. the change in price conditioned on signed trade size, is a key quantity characterizing market liquidity and price dynamics [19, 34]. Besides being of paramount interest for any economic theory of price formation, impact is a major source of transaction costs, which often makes the difference between a trading strategy that is profitable, and one that is not. Hence the interest in this topic is not purely academic in nature.

One of the most surprising empirical finding in the last 25 years is the fact that the impact of a metaorder of total size Q, executed incrementally over time, increases approximately as the square root of Q, and not linearly in Q, as one may have naively expected and as indeed predicted by the now-classic Kyle model [1]. Since impact is non-additive, a natural question concerns the *interaction* of different metaorders executed simultaneously – possibly with different signs and sizes. In particular, one may wonder whether the simultaneous impact of different metaorders could substantially alter the square-root law or conversely whether the square-root law might itself result from the interaction of different metaorders.

Metaorder information is, however, not publicly available, and earlier analyses were mostly based on (often proprietary) data from single financial institutions. These studies give little insight about effects due to the simultaneous execution of metaorders from different investors, which we will call *co-impact* hereafter. Indeed, even if investors individually decide about their metaorders, they might do so based on the same trading signal. Prices can thus be affected by emergent effects such as *crowding*. What is the right way to model the total market impact of simultaneous metaorders on the same asset on the same day? In order to answer this question, we will use a rich dataset concerning the execution of metaorders issued by a heterogeneous set of investors.

This chapter is organized as follows. In Section 5.2 we briefly recall the ANcerno dataset that we used empirically. In Section 5.3 we discuss the limits of the validity of the square-root law on the daily level. In Section 5.4 we find that the market impact of simultaneous daily metaorders is proportional to the square root of their net order flow: this means that the market does not distinguish the different individual metaorders. We then construct a theoretical framework to understand the impact of correlated metaorders in Section 5.5. This allows us to understand when a single asset manager will observe a square root impact law, and when *crowding effects* will lead to deviations from such a behaviour. We also compare in Section 5.5 the results of our simple mathematical model with empirical data, with very satisfactory results. Finally Section 5.6 concludes.

5.2 Data and definitions

Our analysis relies on a rich and heterogeneous database made available by ANcerno, a leading transaction-cost analysis provider[1]. One of the principal advantages of working with such institutional data is that one can simultaneously analyze the trading of many heterogeneous investors executing large orders on the same day and on the same asset. The main caveat though is that one has little knowledge about the motives and style behind the observed portfolio transitions. For example a given metaorder can be part of a longer execution over multiple days. Another possibility is that the final investor may decide to stop a metaorder execution midway if prices move unfavourably. Such effects can potentially bias our results, but we believe that they do not change the qualitative conclusions below.

In the following we will define as a metaorder a series of jointly reported executions performed by a single investor, through a single broker within a single day, on a given stock and in a given direction (buy/sell). However, contrarily to the version of the database used in [10], available labels do not allow us to relate different metaorders executed on behalf of the same final investor by the same or different brokers during the same day. These should ideally be counted as a single metaorder. We will comment later on the biases induced by such a lack of information. Thus each metaorder is characterized by a broker label, the stock symbol, the total volume of the metaorder $|Q|$ and its sign $\epsilon = \pm 1$, and the start time t_{s} and the end time t_{e} of the execution[2].

Our dataset includes the period January 2007 – June 2010 for a total of 880 trading days. Following the filtering procedure introduced in [10] (see Chapter 3 for details) we retain around ~ 8 million metaorders representing around the 5 % of the total reported market volume independently of the year and of the stock capitalization. Without the filters, this number would rise to about 10%. Although we believe that the 90% of missing volume is not *noise*, the ANcerno database is representative of the full set of metaorders. In fact, many of the results discussed are quantitatively similar to those observed on much smaller datasets, such as the CFM database of executed metaorders. In particular, we checked that the results discussed in the present chapter are still valid independently of

[1]See Chapter 3 for more details on the ANcerno dataset.

[2]In this chapter we refer with Q to the *signed* metaorder volume which implies that $\epsilon = \mathrm{sign}(Q) = \pm 1$ (buy/sell).

the use of Filter 3 and Filter 4 introduced in [10]. A particularly important statistic for the following analyses is the number N of simultaneous metaorders in the database, executed on the same stock during the same day. The probability distribution $p(N)$ is shown in the left panel Figure 5.1, indicating that N is broadly distributed with an average close to 5.

The right panel of Figure 5.1 shows the probability distribution of the absolute value of the volume fraction ϕ of the metaorders. This variable plays a key role in the following and is defined as $\phi := Q/V_{\mathrm{d}}$ where V_{d} is the total volume traded during that day. The figure shows that the volume fraction distribution is independent of the metaorder direction (buy or sell) and is also very broad.

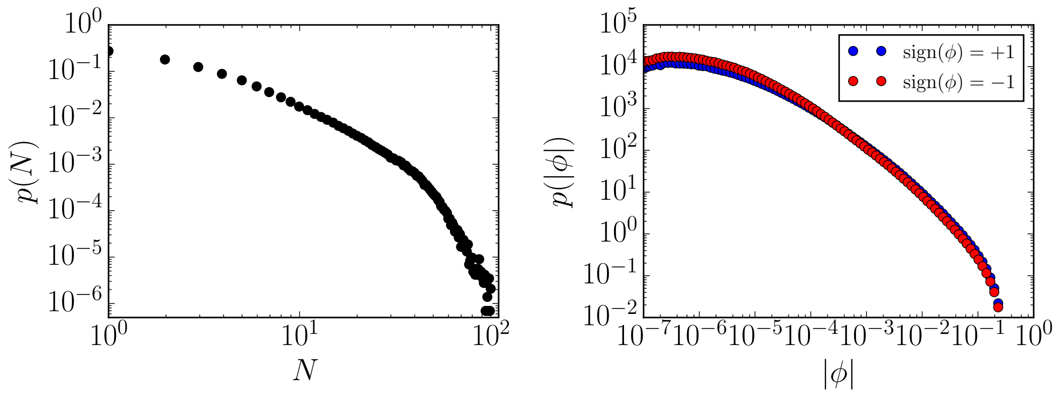

Figure 5.1: (Left panel) Empirical probability distribution of the number N of daily metaorders per asset. (Right panel) Empirical probability distribution of the absolute value of the volume fraction ϕ per metaorder, separately for signs $\epsilon = \mathrm{sign}(\phi) = \pm 1$, i.e. buy/sell.

5.3 The domain of validity of the square-root law

We will quantify market impact in terms of the rescaled log-price $s = \log(S)/\sigma_{\mathrm{d}}$, where S is the market mid-price, which we normalize by the daily volatility of the asset defined as $\sigma_{\mathrm{d}} = (S_{\mathrm{high}} - S_{\mathrm{low}})/S_{\mathrm{open}}$ based on the daily high, low and open prices. In this chapter we will define impact as the expected change of s between the open and the close of the day. This choice will avoid an elaborate analysis of when precisely each metaorder starts and ends, how they overlap and which reference prices to take in each case[3]. When a metaorder of total volume Q is executed, its impact will be defined as

$$I(\phi) := \mathbb{E}[s_{\mathrm{close}} - s_{\mathrm{open}}|\phi], \tag{5.1}$$

[3]To note that in this chapter we consider for each metaorder the daily price change independently of its execution duration $T = t_{\mathrm{e}} - t_{\mathrm{s}}$

for a given metaorder signed volume fraction ϕ.

Empirically, impact is found to be an odd function of ϕ, displaying a concave behavior in $|\phi|$. It is well described by the square root law [2, 4, 5, 9, 12, 46, 50, 82]

$$I(\phi) = Y \times \phi^{\bullet\delta}, \tag{5.2}$$

where here and throughout the chapter we will denote the sign-power operation by $x^{\bullet\delta} := \text{sign}(x) \times |x|^{\delta}$. The dimensionless coefficient Y (called the *Y-ratio*) is of order unity and the exponent δ is in the range 0.4–0.7. It is interesting to note that in Eq. (5.2) only the volume fraction ϕ matters, the time taken to complete execution or the presence of other active metaorders is not directly relevant (remember that the volatility of the instrument has been subsumed in the definition of the rescaled price s). This formula is surprisingly universal across financial products, market venues, time periods and the strategies used for execution.

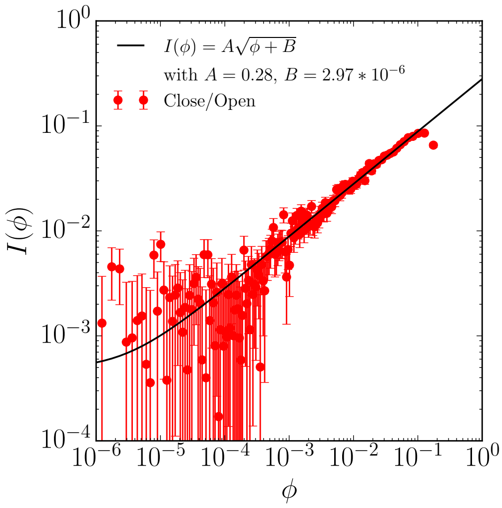

Figure 5.2: Market impact curve $I(\phi) = \mathbb{E}[s_{\text{close}} - s_{\text{open}}|\phi]$ as a function of the metaorder size ratio $\phi = Q/V_{\text{d}}$ computed using the filtered metaorders from the ANcerno dataset in the period from January 2007 – June 2010. We also show the simple fit $I(\phi) = A\sqrt{\phi + B}$, with $A = 0.28$ and $B = 2.97{*}10^{-6}$, which captures some – but not all – of the discrepancy with the square root law at small ϕ.

We first check this empirical result on our dataset. In Figure 5.2 we show the market impact curve obtained by dividing the data into evenly populated bins according to the volume fraction ϕ and computing the conditional expectation of impact for each bin. Here and in the following of the chapter, error bars are determined as standard errors. Note that in all the following empirical plots the price impact curves are normalized by their *Y-ratio* and we will abuse the notation $I(\phi)$ in order to denote the symmetrized measure $I(\phi) = \mathbb{E}[\epsilon \cdot (s_{\text{close}} - s_{\text{open}})||\phi|]$, with $\epsilon = \text{sign}(\phi)$, due to the antisymmetric nature of $I(\phi)$.

While the square-root law holds relatively well when $10^{-3} \lesssim \phi \lesssim 10^{-1}$, three other regimes seem to be present:

1. For very small volume fractions up to $\phi \lesssim 10^{-4}$, impact appears to saturate to a finite, positive value.

2. In the intermediate regime $10^{-4} \lesssim \phi \lesssim 10^{-3}$, impact is closer to a linear function, although the data is very noisy.

3. In the large ϕ regime $\phi \gtrsim 10^{-1}$, impact seems to saturate, or even to decrease with increasing ϕ.

These results are robust across time periods and market capitalizations, and consistent with Ref. [10], where regimes 2. and 3. were also clearly observed. In the following, we will discard altogether the last, large ϕ regime, which is most likely affected by conditioning effects (for example buying more when the price moves down and less when the price moves up). We will on the other hand seek to understand the other three regimes within a consistent mathematical framework.

Intuitively, the breakdown of the square-root law for small ϕ comes from the fact that the signs of the metaorders in our dataset are correlated – particularly so because some metaorders are originating from the same final investor. Let us illustrate the effect of correlations on a simplistic example: imagine that simultaneously to the considered buy metaorder (with volume fraction $\phi > 0$), another metaorder with the same sign and volume fraction $\phi_m > 0$ is also traded. Assuming that the square-root law applies for the combined metaorder (a hypothesis that we will confirm on data), the observed impact should read

$$I(\phi + \phi_m) = Y \times \sqrt{\phi + \phi_m}. \tag{5.3}$$

This tends to the value $Y\sqrt{\phi_m}$ when $\phi \to 0$, behaves linearly when $\phi \ll \phi_m$ and as a square root when $\phi \gg \phi_m$. We show in Figure 5.2 that this simple fit captures some, but not all, of the discrepancy with the square-root law at small ϕ. In particular the intermediate linear region is not well accounted for. We will develop in the following a mathematical model that reproduces all these effects.

A way to minimize the effect of correlations is to restrict to days/assets where there is a unique metaorder in the dataset ($N = 1$). As shown in Figure 5.3, impact in this case is almost perfectly fitted by a square-root law. Figure 5.3, also shows that as N increases, significant departures from the square root law can be observed for small ϕ, as suggested by our simple model Eq. (5.3). An alert reader may however object that the ANcerno database represents a small fraction ($\sim 5\%$) of the total volume. Even when a single metaorder is reported, many other metaorders are likely to be simultaneously present in the market. So why does one observe a square-root law at all, even for single metaorders? The solution to precisely this paradox is one of the main messages of this chapter.

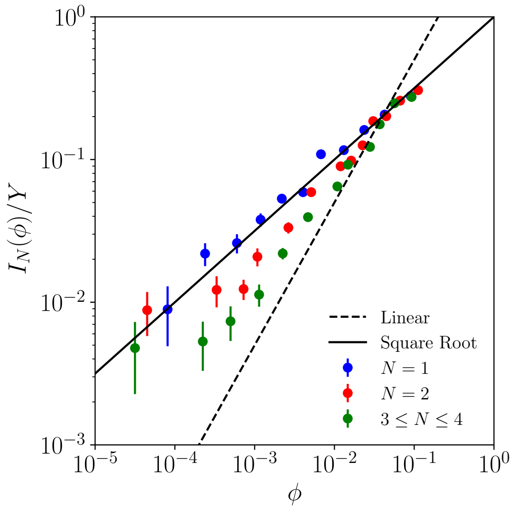

Figure 5.3: Empirical evidence on the effect of the number of metaorders N on the daily price impact curves $I_N(\phi)$ normalized by the prefactor *Y-ratio*: significant departures from the square root law can be observed for small ϕ increasing the number N of daily metaorders per asset.

5.4 How do impacts add up?

In the previous section we showed that the number of metaorders in the market strongly influences how price impact behaves, but we have yet to provide insight into why this is the case. As a first step, we want to determine an explicit functional form of the aggregated market impact of N simultaneous metaorders. As we have emphasized, impact is non-linear, so aggregation is a priori non-trivial. Should one add the square root impact of each metaorder, or should one first add the signed volume fractions before taking the square root? Since orders are anonymous and indistinguishable, the second procedure looks more plausible. This is what we test now. Consider the average aggregate impact conditioned to the co-execution of N metaorders:

$$\mathcal{I}(\boldsymbol{\varphi}_N) = \mathbb{E}[s_{\text{close}} - s_{\text{open}}|\boldsymbol{\varphi}_N], \tag{5.4}$$

where $\boldsymbol{\varphi}_N := (\phi_1, \cdots, \phi_N)$. We make the following parametric ansatz for this quantity:

$$\mathcal{I}(\boldsymbol{\varphi}_N) = Y \times \left(\sum_{i=1}^{N} \phi_i^{\bullet\alpha}\right)^{\bullet\delta/\alpha}, \tag{5.5}$$

where, again, $x^{\bullet\alpha}$ is the signed power of x. By construction this formula is invariant under the permutation of metaorders, as it should be since they are indistiguishable. Y and δ set, respectively, the scale and the exponent of the impact function. The free parameter α interpolates between the case when impacts add up ($\alpha = \delta$) and when only the net traded volume is relevant ($\alpha = 1$).

Figure 5.4 shows the quality of the fit obtained by least squares regressions of Eq. (5.5) for a grid of (α, δ) pairs. We find that the coefficient of determination $r^2(\alpha, \delta)$ of the fit is maximized close to the point $\alpha = 1.0$ and $\delta = 0.5$, which suggests that the aggregated price impact $\mathcal{I}(\boldsymbol{\varphi}_N)$ of N metaorders at the daily scale only depends on the total net order flow, i.e.[4]

$$\mathcal{I}(\boldsymbol{\varphi}_N) \approx Y \times \Phi^{\bullet 1/2}, \tag{5.6}$$

where $\Phi = \sum_{i=1}^{N} \phi_i$. In other words, the market only reacts to the net order flow, not to the way in which this order flow is distributed across investors.

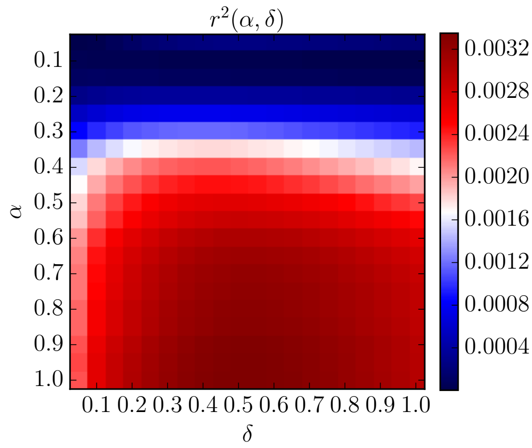

Figure 5.4: The computed coefficient of determination $r^2(\alpha, \delta)$ of the least squares regressions of Eq. (5.5) for a grid of (α, δ) pairs. The coefficient of determination is maximized ($r^2 = 0.0035$) in the vicinity of the point $\alpha = 1.0$ and $\delta = 0.5$.

One can now plot $\mathcal{I}(\Phi)$ as a function of Φ for various N, see Figure 5.5. One clearly sees that provided Φ is not too small, impact is independent of N and crosses over from linear to square root behavior as Φ increases. The linear behavior is in fact more pronounced at smaller values of N. We now turn to a theoretical analysis that will allow us to quantify more precisely the *co-impact* problem, and how the square-root law can survive at large N.

[4]We have in fact tested that the assumption $\alpha = 1$ is also favoured for a general, non parametric shape for the impact function $\mathcal{I}(\Phi)$.

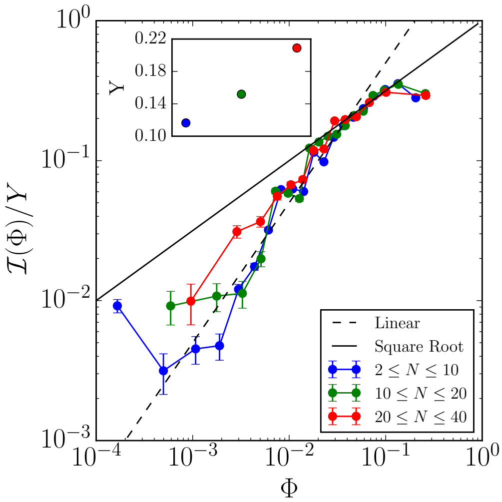

Figure 5.5: Global market impact $\mathcal{I}(\Phi) = \mathbb{E}[\mathcal{I}(\boldsymbol{\varphi}_N)|\sum_{i=1}^N \phi_i = \Phi]$ normalized by the prefactor Y as a function of the net order flow Φ for various buckets in N and symmetrized as explained in Section 5.3. The inset shows the normalization Y for the different curves: the dependence on N is a consequence of the sampling which is removed considering the multiplicative factor $\langle \sigma_\mathrm{d}/\sqrt{V_\mathrm{d}} \rangle_N$ given by the conditional average for a fixed range of N, i.e. $Y \times \langle \sigma_\mathrm{d}/\sqrt{V_\mathrm{d}} \rangle_N \simeq 4.25 \times 10^{-6}$ (not shown).

5.5 Correlated metaorders and co-impact

5.5.1 The mathematical problem

Even if an asset manager knows the average impact formula Eq. (5.6), this may not be sufficient to estimate his actual impact which depends strongly on the presence of other contemporaneous metaorders.

Suppose the manager k wants to execute a volume fraction $\phi_k = \phi$. If all the other $N-1$ metaorders were known, the daily price impact would be given by the global impact function $\mathcal{I}(\Phi)$ determined in the previous section, with $\Phi = \phi_k + \sum_{i \neq k}^N \phi_i$. However, this information is obviously not available to the manager k. His best estimate of the average impact given N is the conditional expectation

$$I_N(\phi) = \mathbb{E}[\mathcal{I}(\Phi)|\phi_k = \phi] = \mathbb{E}\Big[\mathcal{I}\Big(\phi_k + \sum_{i \neq k}^N \phi_i\Big)\Big|\phi_k = \phi\Big] \tag{5.7}$$

over the conditional distribution $P(\boldsymbol{\varphi}_N|\phi_k = \phi)$ of the metaorders. Since the number of metaorders is in general not known either, the expected individual market impact is given by

$$I(\phi) = Y \times \sum_N p(N) \int_{-\infty}^{+\infty} \mathrm{d}\phi_1 \dots \mathrm{d}\phi_N P(\boldsymbol{\varphi}_N|\phi_k = \phi)\Big(\phi_k + \sum_{i\neq k}^{N} \phi_i\Big)^{\bullet 1/2}, \tag{5.8}$$

where we have used Eq. (5.6). In such a way to compute $I_N(\phi)$ and $I(\phi)$ we need to know the joint probability density function $P(\boldsymbol{\varphi}_N) := P(\phi_1, \dots, \phi_N)$, which is in general a complicated and high-dimensional object. Then to create a tractable model that can be calibrated on data, we must make some reasonable assumption on the dependence structure of the ϕ_i. In the next subsection we investigate the simple case where the ϕ_i are all independent, and then turn to an empirical characterization of the correlations between metaorders. We finally provide the results of our empirically inspired model and compare them with the empirical market impact curves.

5.5.2 Independent metaorders

The simplest assumption about the form of $P(\boldsymbol{\varphi}_N)$ is that metaorder volumes are i.i.d., meaning

$$P(\boldsymbol{\varphi}_N) = \prod_{i=1}^{N} p(\phi_i). \tag{5.9}$$

Assuming for simplicity that each ϕ_i is a Gaussian random variable with zero mean and variance Σ_N^2, where the lower index indicates an explicit dependence on N. Thus $N-1$ simultaneous metaorders generate a Gaussian noise contribution of amplitude $\Sigma_N\sqrt{N-1}$ on top of $\phi_k = \phi$. In Appendix 5.A.1 we show analytically that:

- For small metaorders the noise term dominates, leading to

$$I_N(\phi) \propto \phi \quad \text{when} \quad \phi \ll \phi_N^* := \Sigma_N\sqrt{N-1}.$$

- For large metaorders the $N-1$ other simultaneous metaorders can be neglected and thus

$$I_N(\phi) \propto \sqrt{\phi} \quad \text{when} \quad \phi \gg \phi_N^*.$$

In Appendix 5.A.1 we show that the above results remain valid in the limit of large N independently of the shape of the volume distribution provided its variance is finite.

The full analytical solutions for different N values, but *fixed* $\Sigma_N = \Sigma$, are shown in the left panel of Figure 5.6. One clearly sees the cross-over from a linear behavior at small ϕ to a square root at larger ϕ. However, one expects Σ_N to *decrease* with N, simply because as the number of metaorders increases, the volume fraction represented by each of them must decrease[5]. As shown in the right panel of Figure 5.6, this is the case empirically since for $N \gtrsim 10$, Σ_N indeed decays as N^{-1} (as also suggested

[5]For example, the variance of a flat Dirichlet random variable $(X_1, ..., X_N) \sim \mathrm{Dir}(N)$ describing fractions is $\mathbb{V}[X_i|N] = (N-1)/(N^2(N+1)) \sim N^{-2}$.

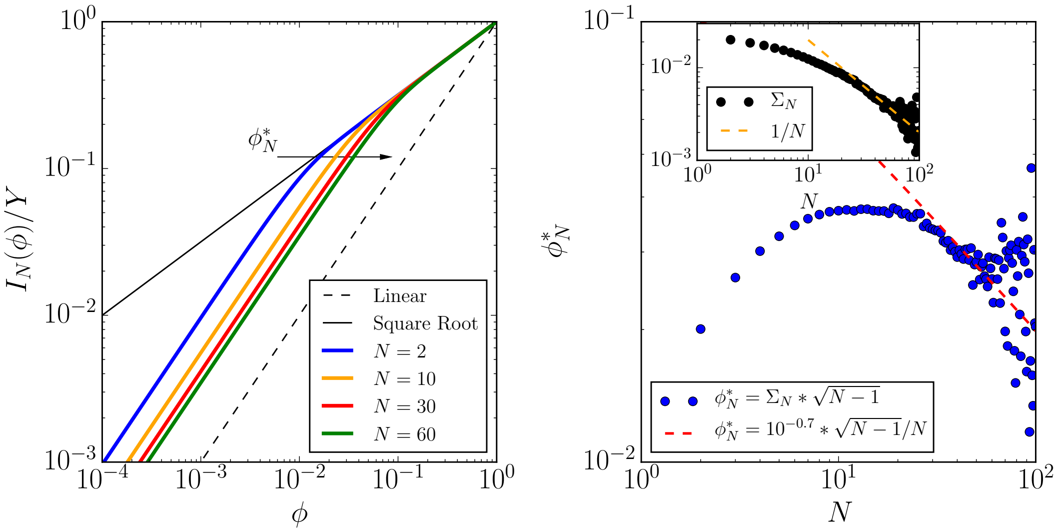

Figure 5.6: (Left panel) Market impact curves $I_N(\phi)/Y$ for i.i.d. Gaussian metaorders for different $N \in \{2, 10, 30, 60\}$ and *fixed* $\Sigma_N = \Sigma = 0.8\%$ computed from empirical data. The transition from the square root to the linear regime takes place for $\phi_N^* \simeq \Sigma\sqrt{N-1}$. (Right panel) Empirically estimated ϕ_N^* as a function of the number N of daily metaorders per asset. This is obtained by computing the empirical order size's standard deviation $\Sigma_N = \sqrt{\mathbb{V}[\phi|N]}$ conditioned to N. The dashed red line shows the case in which $\Sigma_N \sim 1/N$ for each N: in the inset we report the standard deviation Σ_N of the order size ϕ as a function of the number N of daily metaorders per asset.

by Figure 5.7 below). Hence, for large N, the crossover value ϕ_N^* *decreases* with N as $N^{-1/2}$. This explains why the square root law can at all be observed when a large number of metaorders are present. If these metaorders are independent, their net impact on the price averages out, leaving the considered metaorder as if it was alone in a random flow, as assumed in theoretical models [2, 82]. We now turn to the effect of correlations between metaorders.

5.5.3 Metaorder correlations

In order to build a sensible model of $P(\boldsymbol{\varphi}_N)$ we consider separately the size distribution and the size cross-correlations. From the right panel of Figure 5.1 we observe that the marginals $p(\phi_i)$ are to a good approximation independent of the direction, buy or sell, and moderately fat tailed. The latter observation suggests that the total net order flow $\Phi = \sum_{i=1}^{N} \phi_i$ is *not* dominated by a single metaorder. A way to quantify this is through

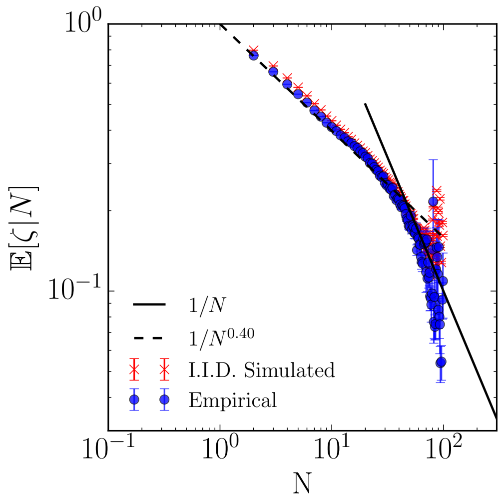

Figure 5.7: Average value of the Herfindahl index $\mathbb{E}[\zeta|N]$ as a function of the number N of daily metaorders per asset computed from the empirical data (blue dot symbols) and simulating i.i.d. $|\phi_i|$ (red star symbols) extracted from the empirical distribution illustrated in the right panel of Figure 5.1.

the Herfindahl index (or inverse participation ratio) ζ, defined as:

$$\zeta := \frac{\sum_{i=1}^{N} \phi_i^2}{\left(\sum_{i=1}^{N} |\phi_i|\right)^2}. \tag{5.10}$$

This quantity is of order $1/N$ if all metaorders are of comparable size, and of order 1 if one metaorder dominates. In Figure 5.7 we show the dependence of $\mathbb{E}[\zeta|N]$ as a function of N, which clearly decays with N. It also compares very well to the result obtained assuming the absolute volume fractions $|\phi_i|$ to be independent, identically distributed variables, drawn according to the empirical distribution shown in Figure 5.1. We therefore conclude that:

- Metaorders in the ANcerno database are typically of comparable relative sizes ϕ.
- Absolute volume correlations do not play a major role, and we will neglect them henceforth.

Sign correlations, on the other hand, *do* play an important role in determining the impact of simultaneous metaorders. The empirical average sign correlation of metaorders simultaneously executed on the same asset is defined as

$$\mathbb{C}_\epsilon(N) := \frac{\mathbb{E}[\epsilon_i \epsilon_j|N] - \mathbb{E}[\epsilon_i|N]^2}{\mathbb{E}[\epsilon_i^2|N] - \mathbb{E}[\epsilon_i|N]^2}\,, \tag{5.11}$$

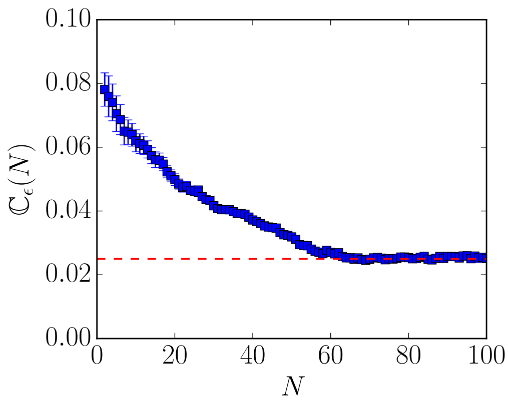

Figure 5.8: Empirical average sign correlation $\mathbb{C}_\epsilon(N)$ as a function of the number N of daily metaorders: the dashed red line represents the plateau value $\mathbb{C}_\epsilon \approx 0.025$ at large N, which, we believe, is a reasonable proxy for the correlation of orders submitted by different asset managers.

where $\mathbb{E}[\cdots|N]$ is the average over all days and assets such that exactly N metaorders were executed. Figure 5.8 shows the dependence of $\mathbb{C}_\epsilon$ on N. We clearly see that on average the daily metaorders executed on the same asset are positively correlated. Furthermore, $\mathbb{C}_\epsilon(N)$ is seen to decrease as N increases. This is likely due to the fact that there are multiple concurrent metaorders submitted by the same manager, an effect that becomes less prominent as N increases. The plateau value $\mathbb{C}_\epsilon \approx 0.025$ at large N is, we believe, a reasonable proxy for the correlation of orders submitted by different asset managers.

5.5.4 Market impact with correlated metaorders

A natural model would be to consider the ϕ_i's as exchangeable multivariate Gaussian variables of zero mean, variance Σ_N^2 and cross-correlation coefficient $\mathbb{C}_\phi(N)$. Appendix 5.A.2 shows that the qualitative behavior for independent metaorders remains the same when $\mathbb{C}_\phi(N) > 0$. Specifically, one finds that the average impact $I_N(\phi)$ can be obtained by making the substitution

$$\phi \to \phi[1 + (N-1)\mathbb{C}_\phi(N)] \tag{5.12}$$

in the expression of $I_N(\phi)$ for independent Gaussians. This is expected, as $(N-1)\mathbb{C}_\phi(N)$ gives the effective number of additional volume-weighted metaorders correlated to the original one. By the same token though, $I_N(\phi)$ still vanishes linearly for small ϕ, whereas empirical data suggests a positive intercept when $\phi \to 0$.

As an alternative model that emphasizes sign-correlations, let us assume that the joint distribution of the ϕ_i's can be written as

$$P(\boldsymbol{\varphi}_N) = \mathcal{P}(\boldsymbol{\epsilon}_N) \prod_{i=1}^{N} p(|\phi_i|), \tag{5.13}$$

meaning that metaorder sizes are independent, while the signs are possibly correlated. This specific form is motivated by the observation that the size of a metaorder is mainly related to the assets under management of the corresponding financial institution, while the sign is related to the trading signal. One can expect that different investors use correlated information sources, while the size of the trades is idiosyncratic.

We further assume that there is a unique common factor determining the sign of the metaorders. In other words, the statistical model for the signs is the following:

$$\mathbb{P}(\epsilon_i = +1|\tilde{\epsilon}) = \frac{1}{2}(1 + \gamma_\epsilon \tilde{\epsilon}); \qquad \mathbb{P}(\epsilon_i = -1|\tilde{\epsilon}) = \frac{1}{2}(1 - \gamma_\epsilon \tilde{\epsilon}), \tag{5.14}$$

where $\tilde{\epsilon}$ is the hidden sign factor, such that $\mathbb{P}(\tilde{\epsilon} = \pm 1) = 1/2$, and γ_ϵ is the sign correlation between each sign ϵ_i and the hidden sign factor $\tilde{\epsilon}$. A simple calculation leads to

$$\mathbb{C}_\epsilon(N) = \mathbb{P}(\epsilon_i = \epsilon_j) - \mathbb{P}(\epsilon_i = -\epsilon_j) = \gamma_\epsilon^2 \tag{5.15}$$

where we omitted the γ_ϵ's explicit dependence on N implied by the sign correlation $\mathbb{C}_\epsilon(N)$. Contrarily to the Gaussian case, we have not been able to obtain analytical formulas, but instead relied on numerical simulations to obtain $I_N(\phi)$ for different combinations of $\mathbb{C}_\epsilon(N)$ and N, reported in Figure 5.9. Results for unsigned volumes generated from a half-normal distribution calibrated on data are shown in Figure 5.9. We observe that the individual price impact $I_N(\phi)$ converges to a positive constant $I_N(0) > 0$ when $\phi \to 0$, despite $I_N(0) = 0$ for a Gaussian model. For intermediate ϕ, $I_N(\phi)$ is linear and it crosses over at larger ϕ to a square root. For fixed N the intercept value increases with the sign correlation $\mathbb{C}_\epsilon$, see the right panel of Figure 5.9. The intuition is that conditioned to the fact that I buy, and independently of the size of my trade, the order flow of other actors will be biased towards buy as well, and I will suffer from the impact of their trades. In fact, subtracting the non-zero intercept of $I_N(\phi)$ leads to impact curves that look almost identical to those of Figure 5.6, i.e. a linear region for small ϕ followed by a square root region beyond a crossover value $\phi_N^* \sim \Sigma_N \sqrt{N-1}$. Since for large N $\phi_N^* \to 0$, one simply expects a square-root law, shifted by the intercept $I_N(0)$.

5.5.5 Empirical calibration of the model

With the aim to compare the model prediction with empirical data, we propose a calibration method described in Appendix 5.A.4. This is based on the assumption that metaorder signs are independent random variables sampled from a half-Gaussian calibrated on empirical data. The metaorder sign correlation structure can be estimated by introducing a realized sign correlation

$$\rho_\epsilon := \frac{2}{N(N-1)} \sum_{1 \le i < j \le N} \epsilon_i \epsilon_j, \tag{5.16}$$

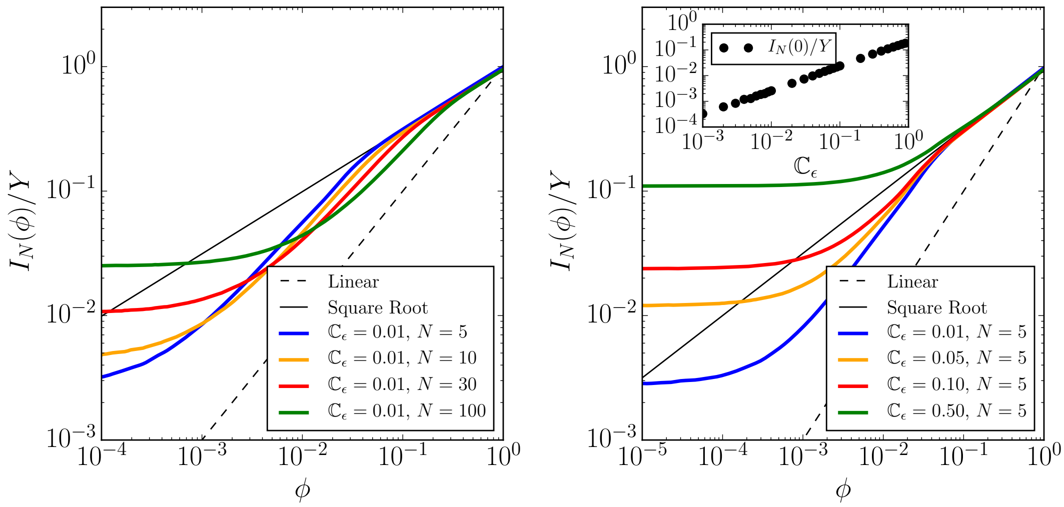

Figure 5.9: Market impact curves $I_N(\phi)/Y$ computed for the correlated signs model, where volumes are drawn according to a half-normal distribution with $\mathbb{E}[|\phi_i|] = \Sigma\sqrt{2/\pi}$ and $\Sigma = 0.8\%$ is set equal to its empirical value averaged over N. (Left panel) Numerical simulations for fixed sign correlation $\mathbb{C}_\epsilon = 0.01$ but varying number of metaorders $N \in \{5, 10, 30, 100\}$. (Right panel) Numerical simulations for fixed number of metaorders $N = 5$ but with varying sign correlation $\mathbb{C}_\epsilon \in \{0.01, 0.05, 0.1, 0.5\}$: as shown in the inset the intercept $I_N(0)/Y$ decreases linearly with the sign correlation for $\mathbb{C}_\epsilon \to 0$. Note that the individual price impact $I_N(\phi)$ values converges to a positive constant when $\phi \to 0$. For intermediate ϕ, $I_N(\phi)$ is linear and crosses-over at larger ϕ to a square root.

which is then used to estimate the sign correlation $\mathbb{C}_\epsilon(N)$ of Eq. (5.15). Once the model is calibrated, we use numerical simulations to compute the expected market impact $I(\phi)$, see Appendix 5.A.4 for the precise details of the procedure.

Figure 5.10 shows that imposing correlation only between the signs leads to a very good prediction of the empirical curves, justifying the adoption of the sign correlated model. All the features of the empirical impact curves are qualitatively well reproduced, in line with Figure 5.2. This includes the clear deviations from the square root law for $\phi \leq 10^{-3}$ with both a linear regime and a constant price impact I_0 when $\phi \to 0$.

5.6 Conclusions

It is a commonly acknowledged fact that market prices move during the execution of a trade – they increase (on average) for a buy order and decrease (on average) for a sell order. This is, loosely stated, the phenomenon known as market impact. In this chapter we have presented one of the first studies breaking down market impact of metaorders

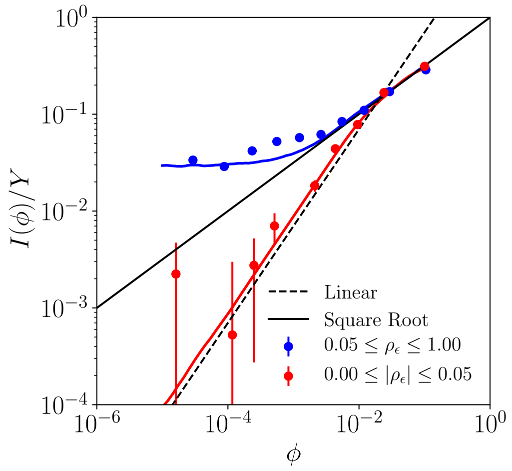

Figure 5.10: Comparison between calibrated sign-correlated model (colored lines) and empirical data (circles) for the case $N \in [2, 10]$: the sample is split into two sub-samples respectively with $\rho_\epsilon \geq 0.05$ and $0 \leq |\rho_\epsilon| \leq 0.05$. The theoretical curves are calculated through numerical simulations as explained in the main text (see Appendix 5.A.4 for all the details of the procedure).

executed by different investors, and taking into account *interaction/correlation* effects. We investigated how to aggregate the impact of individual actors in order to best explain the daily price moves. The large number and heterogeneity of the metaorders traded by financial institutions allows precise measurements of price impact in different conditions with reduced uncertainty. We found that both the number of actors simultaneously trading on a stock and the *crowdedness* of their trade (measured by the correlation of metaorder signs) are important factors determining the impact of a given metaorder.

Our main conclusions are as follows:

- The market chiefly reacts to the total net order flow of ongoing metaorders, the functional form being well approximated by a square root at least in a range of volume fraction ϕ. As expected in anonymous markets, impact is insensitive to the way order flow is distributed across different investors.

- The number N of executed metaorders and their mutual sign correlations $\mathbb{C}_\epsilon$ are relevant parameters when an investor wants to precisely estimate the market impact of their own metaorders.

- Using a simple heuristic model calibrated on data, we are able to reproduce to a good level of precision the different regimes of the empirical market impact curves, as a function of ϕ, N and $\mathbb{C}_\epsilon$.

- When the number of metaorders is not large, and when $\mathbb{C}_\epsilon > 0$, a small investor will observe *linear* impact with a non-zero intercept I_0, crossing over to a square-root law at larger ϕ. The intercept I_0 grows with $\mathbb{C}_\epsilon$ and can be interpreted as the average impact of all other metaorders.
- When the number of metaorders is very large and the investor has no correlation with their average sign, they should expect on a given day a square root impact randomly shifted upwards or downwards by I_0. Averaged over all days, a pure square-root law emerges, which explains why such behavior has been reported in many empirical papers.

On the last point, we believe that our study sheds light on an apparent paradox: How can a non-linear impact law survive in the presence of a large number of simultaneously executed metaorders? As we have seen, the reason is that for a metaorder uncorrelated with the rest of the market, the impacts of other metaorders cancel out on average. Conversely, any intercept of the impact law can be interpreted as a non-zero correlation with the rest of the market. Given the importance of the subject, our results present several interesting applications. Our aggregated price impact model should be of interest both to practitioners trying to monitor and reduce their trading costs, and also to regulators that seek to improve the stability of markets.

Appendix

5.A From the bare to the market impact function

The expected individual market impact of a metaorder with signed volume ϕ is estimated by

$$I(\phi) = \sum_N p(N) I_N(\phi) \tag{5.17}$$

where $p(N)$ is the probability distribution of the daily number N of metaorders per asset and

$$I_N(\phi) = \mathbb{E}[\mathcal{I}(\boldsymbol{\varphi}_N)|\phi_k = \phi] = \int_{-\infty}^{\infty} \mathrm{d}\phi_1 \cdots \mathrm{d}\phi_N P(\boldsymbol{\varphi}_N|\phi_k = \phi) \left(\phi_k + \sum_{i \neq k}^{N} \phi_i\right)^{\bullet 1/2} \tag{5.18}$$

is the market impact computed from the bare impact function $\mathcal{I}(\boldsymbol{\varphi}_N) := (\sum_{i=1}^{N} \phi_i)^{\bullet 1/2}$ with fixed N. Assuming that $p(N)$ is kwown a priori the market impact $I_N(\phi)$ is given from the expectation in Eq. (5.18) done over the conditional probability distribution $P(\boldsymbol{\varphi}_N|\phi_k = \phi)$ of the volume metaorders. However, in such a way to do analytical computation it is necessary to assume reasonable hypothesis for the joint distribution function $P(\boldsymbol{\varphi}_N) := P(\phi_1, \cdots, \phi_N)$.

For this reason, we start considering the case of i.i.d. metaorders and we firstly show analytically how the transition from the square root to a linear market impact is possible in the i.i.d. Gaussian framework. Secondly, we generalize these results in the limit of large N for any symmetric volume distribution which sastifies the Central Limit Theorem assumptions. Thirdly, we show that the same results continue to be valid introducing correlation between the signed volumes in a Gaussian framework. For last, we describe how to compute numerically the market impact in Eq. (5.18) in the case of metaorder signs correlated and i.i.d. unsigned volumes.

5.A.1 Market impact with i.i.d. metaorders

In the case of i.i.d. signed metaorders, the conditional joint distribution factorizes as

$$P(\boldsymbol{\varphi}_N|\phi_k = \phi) = \prod_{i\neq k}^{N} p(\phi_i); \tag{5.19}$$

this implies that the price impact $I_N(\phi)$ of a single metaorder $\phi_k = \phi$ out of N is given by

$$I_N(\phi) = \int_{-\infty}^{\infty} \mathrm{d}\phi_m \underbrace{\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\lambda e^{-i\lambda\phi_m}\hat{p}(\lambda)^{N-1}}_{p(\phi_m)}(\phi+\phi_m)^{\bullet 1/2} = \int_{-\infty}^{\infty} \mathrm{d}\phi_m p(\phi_m)(\phi+\phi_m)^{\bullet 1/2} \tag{5.20}$$

where $\phi_m = \sum_{i\neq k}^{N} \phi_i$ is the net order flow executed simultaneously to the metaorder ϕ_k and $\hat{p}(\lambda) = \mathbb{E}[e^{i\lambda\phi_i}]$ is the characteristic function of the signed volume distribution $p(\phi_i)$. Although the introduction of the characteristic function $\hat{p}(\lambda)$ in Eq. (5.20) will be a convenient way to exploit the convergence of the net order flow distribution $p(\phi_m)$ as discussed in the following, the computation of the market impact $I_N(\phi)$ in a analytical closed form is possible only in the Gaussian case.

Independent Gaussian metaorders

In the Gaussian case, i.e. $p(\phi_i) \sim \mathcal{N}(0, \Sigma_N^2)$ with $\Sigma_N^2 = \mathbb{V}[\phi_i|N]$, we can go further analytically in Eq. (5.20) since $\hat{p}(\lambda) = e^{-\lambda^2\Sigma_N^2/2}$. In fact the integral representation of the price impact

$$\begin{aligned} I_N(\phi) &= \frac{1}{\Sigma_N\sqrt{2\pi(N-1)}}\int_{-\infty}^{\infty} \mathrm{d}\phi_m e^{-\phi_m^2/(2\Sigma_N^2(N-1))}(\phi+\phi_m)^{\bullet 1/2} = \\ &= \frac{1}{\Sigma_N\sqrt{2\pi(N-1)}}\int_0^{\infty} \mathrm{d}x\sqrt{x}\Big[e^{-(x-\phi)^2/(2\Sigma_N^2(N-1))} - e^{-(x+\phi)^2/(2\Sigma_N^2(N-1))}\Big] \end{aligned} \tag{5.21}$$

can be expressed in the following analytical way

$$I_N(\phi) = \frac{\Gamma(1/4)}{2\sqrt{\pi}}\frac{\phi}{(2(N-1)\Sigma_N^2)^{1/4}}\, e^{-\frac{\phi^2}{2(N-1)\Sigma_N^2}}\, {}_1F_1\left(\frac{5}{4}, \frac{3}{2}, \frac{\phi^2}{2(N-1)\Sigma_N^2}\right) \tag{5.22}$$

where $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} \mathrm{d}x$ is the Gamma function and

$$ {}_1F_1\left(\frac{5}{4}, \frac{3}{2}, z\right) = \frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{5}{4})} \sum_{j=0}^{\infty} \frac{\Gamma(\frac{5}{4}+j)}{\Gamma(\frac{3}{2}+j)} \frac{z^j}{j!} \tag{5.23} $$

is the Kummer confluent hypergeometric function with $z = \frac{\phi^2}{(2(N-1)\Sigma_N^2)}$ [84].

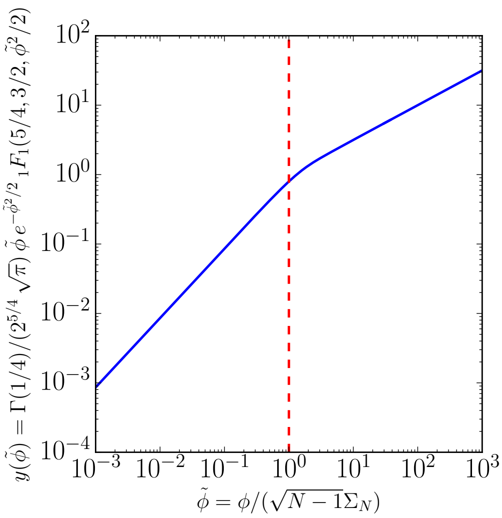

Figure 5.11: Rescaled market impact function $y(\tilde{\phi}) = I_N(\phi)/((N-1)\Sigma_N^2)^{1/4}$ in function of the dimensionless parameter $\tilde{\phi} := \phi/(\sqrt{N-1}\Sigma_N)$ with $I_N(\phi)$ given by Eq. (5.22): the vertical dashed red line represents the transition from the linear impact (left side) to the square root one (right side).

The price impact $I_N(\phi)$ in Eq. (5.22) is shown in the left panel of Figure 5.6 for different N and the parameter Σ_N *fixed*. If the metaorder volume ϕ is smaller than the sum of the other $N-1$ metaorders, i.e. $\phi \ll \phi_m$, then price impact is linear. Instead, when our metaorder dominates, i.e. $\phi \gg \phi_m$, the price impact follows a square root function. The transition from the linear to the square root regime takes place around $\phi_N^* \simeq \Sigma_N\sqrt{N-1}$, where $\Sigma_N\sqrt{N-1}$ is naturally interpreted as a measure for the market noise, in agreement with the change of the functional shape of the rescaled market impact function $y(\tilde{\phi}) = I_N(\phi)/((N-1)\Sigma_N^2)^{1/4}$ represented in Figure 5.11 in function of the adimensional parameter $\tilde{\phi} := \phi/(\sqrt{N-1}\Sigma_N)$. To note furthermore that the linear regime comes out immediately from the expansion of $I_N(\phi)$ in Eq. (5.22) around $\phi = 0$ as follows

$$I_N(\phi) = \frac{1}{\Sigma_N\sqrt{2\pi(N-1)}}\int_0^\infty \mathrm{d}x\sqrt{x}e^{-x^2/(2\Sigma_N^2(N-1))}\Big[e^{-(\phi^2-2x\phi)/(2\Sigma_N^2(N-1))}+$$
$$- e^{-(\phi^2+2x\phi)/(2\Sigma_N^2(N-1))}\Big] \simeq$$
$$\simeq \sqrt{\frac{2}{\pi}}\frac{\phi}{(\Sigma_N^2(N-1))^{3/2}}\int_0^\infty \mathrm{d}x x^{3/2}e^{-x^2/(2\Sigma_N^2(N-1))} = 2^{3/4}\frac{\Gamma(5/4)}{(\pi^2\Sigma_N^2(N-1))^{1/4}}\phi. \quad (5.24)$$

Remark 1. It follows that for i.i.d. Gaussian metaorders the slope of the linear price impact region decreases with N (as shown explicitly in Eq. (5.24)) and the crossover to the square root region happens in ϕ_N^* obtained by solving

$$\xi\frac{\phi_N^*}{(\Sigma_N^2(N-1))^{1/4}} \simeq (\phi_N^*)^{1/2}, \quad (5.25)$$

i.e. $\phi_N^* \simeq \xi^{-1}\Sigma_N\sqrt{N-1}$ with $\xi = 2^{3/4}\Gamma(5/4)/\sqrt{\pi}$.

Limit of large N for generally distributed independent metaorders

The previous conclusions discussed in the i.i.d. Gaussian framework are also valid for other i.i.d. volume distributions as discussed in this Appendix: in fact, we can generalize them in the limit of large N for any symmetric volume distribution $p(\phi_i)$ which sastifies the Central Limit Theorem assumptions.

In the limit of large N, for which the Central Limit Theorem applies if certain conditions (discussed below) on $p(\phi_i)$ are sastisfied, the symmetric volume distribution $p(\phi_m)$ introduced in Eq. (5.20) converges to a stable law $\mathcal{G}_\alpha$ described by a characteristic function

$$\hat{p}(\lambda) = \mathbb{E}[e^{i\lambda\phi_m}] = e^{-c|\lambda|^\alpha} \quad (5.26)$$

where $c \in (0,\infty)$ is the scale parameter and $\alpha \in (0,2]$ is the stability exponent. In other words we say that the volume distribution $p(\phi_i)$ belongs to the *domain of attraction* of the stable distribution $\mathcal{G}_\alpha$ if there exist constants $a_m \in \mathbb{R}$, $b_m > 0$ such that

$$b_m^{-1}(\phi_m - a_m) \to \mathcal{G}_\alpha, \quad (5.27)$$

i.e. the renormalized and recentred sum $\phi_m = \sum_{i\neq k}^N \phi_i$ converges in distribution to $\mathcal{G}_\alpha$. The Central Limit Theorem gives the conditions such that this convergence in distribution to a stable law $\mathcal{G}_\alpha$ is guaranted:

- $p(\phi_m)$ converges to a Gaussian distribution ($\alpha = 2$ in Eq. (5.26)) if and only if

$$\int_{|\phi_i|\le x}\phi_i^2 p(\phi_i)\mathrm{d}\phi_i \quad (5.28)$$

is a slowly varying function $L(x)$, i.e. $\lim_{x\to\infty} L(tx)/L(x) = 1$ for all $t > 0$; then it follows that if $\Sigma_N^2 = \mathbb{V}[\phi_i|N] < \infty$

$$((N-1)^{1/2}\Sigma_N)^{-1}\phi_m \longrightarrow \mathcal{N}(0,1), \tag{5.29}$$

while if $\mathbb{V}[\phi_i|N] = \infty$

$$((N-1)^{1/2}L_1)^{-1}\phi_m \longrightarrow \mathcal{N}(0,1) \tag{5.30}$$

with L_1 a slowly varying function and $\mathcal{N}(0,1)$ a Gaussian distribution with mean zero and variance 1.

- $p(\phi_m)$ converges to a Lévy distribution (for some $\alpha < 2$ in Eq. (5.26)) if and only if

$$\int_{-\infty}^{-x} p(\phi_i)\mathrm{d}\phi_i = \frac{c_1 + o(1)}{x^\alpha}L(x), \quad 1 - \int_{-\infty}^{x} p(\phi_i)\mathrm{d}\phi_i = \frac{c_2 + o(1)}{x^\alpha}L(x), \quad x \to \infty \tag{5.31}$$

where $L(x)$ is a slowly varying function and c_1, c_2 are non-negative constants such that $c_1 + c_2 > 0$; then it follows that

$$((N-1)^{1/\alpha}L_2)^{-1}\phi_m \longrightarrow \mathcal{L}(c,\alpha) \tag{5.32}$$

with L_2 a slowly varying function and $\mathcal{L}(c,\alpha)$ a Lévy distribution with scale parameter $c \in (0,\infty)$ and stability exponent $\alpha \in (0,2)$.

Remark 2. It follows immediately that for any volume distributions $p(\phi_i)$ belonging to the *domain of attraction* of the normal law, i.e. satisfying the condition in Eq. (5.28), the price impact $I_N(\phi)$ in the limit of large N is described by Eq. (5.22) and the transition from a linear to a square root price impact discussed in the Remark 1 continue to be valid: to note that in the case of $\mathbb{V}[\phi_i|N] = \infty$ it is sufficient to substitute Σ_N in Eq. (5.25) with the appropriate slowly varying function L_1.

Moreover we can show that the transition from a linear to a square root regime is still present for volume distributions $p(\phi_i)$ belonging to the *domain of attraction* of the Lévy distribution, i.e. that sastify the condition in Eq. (5.31) with $\alpha < 2$ and then $p(\phi_m) \to \mathcal{L}(c(N-1)^{1/\alpha}, \alpha)$. In fact, expanding to first order the bare impact function $\mathcal{I}(\phi, \phi_m) = (\phi + \phi_m)^{\bullet 1/2}$ around $\phi = 0$ in Eq. (5.20)

$$I_N(\phi) \simeq \frac{\phi}{2}\int_{-\infty}^{\infty} \mathrm{d}\phi_m p(\phi_m)|\phi_m|^{-1/2} = \frac{\phi}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}\lambda \hat{p}(\lambda)\int_{-\infty}^{\infty} \mathrm{d}\phi_m e^{-i\lambda\phi_m}|\phi_m|^{-1/2}, \tag{5.33}$$

introducing the characteristic function of the net order flow ϕ_m given by

$$\hat{p}(\lambda) = e^{-c(N-1)^{1/\alpha}|\lambda|^\alpha} \tag{5.34}$$

for large N and taking present that the last integral in Eq. (5.33) is a known Fourier transform

$$\int_{-\infty}^{+\infty} \mathrm{d}\phi_m e^{-i\lambda\phi_m}|\phi_m|^{-1/2} = \frac{\sqrt{2\pi}}{|\lambda|^{1/2}} \tag{5.35}$$

we obtain a linear price impact

$$I_N(\phi) \simeq \frac{\phi}{2\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{p}(\lambda)|\lambda|^{-1/2}\mathrm{d}\lambda. \tag{5.36}$$

Though it is not possible to analytically compute the last integral, its behavior for large N is known using the saddle point approximation or Perron's method: since all the conditions of the theorem at page 105 of Ref. [85] are satisfied, one can approximate the integral in Eq. (5.36) as follows

$$\int_{-\infty}^{+\infty} \hat{p}(\lambda)|\lambda|^{-1/2}\mathrm{d}\lambda \sim \Gamma\left(\frac{1}{2\alpha}\right) \frac{1}{\alpha c^{1/(2\alpha)}(N-1)^{1/(2\alpha)}}. \tag{5.37}$$

Remark 3. In the limit of large N it is then possible to show analytically that for volume distributions $p(\phi_i)$ belonging to the *domain of attraction* of the Lévy distribution the price impact is characterized by a linear regime described by

$$I_N(\phi) \simeq \frac{1}{2\sqrt{2\pi}}\Gamma\left(\frac{1}{2\alpha}\right) \frac{\phi}{\alpha[c(N-1)]^{1/(2\alpha)}}, \tag{5.38}$$

followed by a transition to a square root one around $\phi_N^* \simeq (c(N-1))^{1/\alpha}$.

5.A.2 Market impact with correlated Gaussian metaorders

With the aim to introduce correlations between the metaorder volumes $\boldsymbol{\varphi}_N = (\phi_1, \cdots, \phi_N)$ in the Gaussian framework it is useful to define the following joint probability distribution

$$P(\boldsymbol{\varphi}_N) = \frac{1}{Z_N} \exp\left(-\frac{A_N}{2}\sum_{i=1}^N \phi_i^2 + \frac{B_N}{N}\sum_{i<j}^N \phi_i\phi_j + \mu\sum_{i=1}^N \phi_i\right), \tag{5.39}$$

where Z_N is a normalization function, A_N and B_N are parameters depending on N and μ is an external field.

Calibration from data: means and correlations

The first step to calibrate $P(\boldsymbol{\varphi}_N)$ in Eq. (5.39) is to express the model parameters A_N, B_N and μ in terms of observable quantities, namely $\mathbb{E}[\phi_i\phi_j|N]$ and $\mathbb{E}[\phi_i|N]$. Due to the presence of an interaction term, the computation of Z_N requires the use of a Hubbard-Stratonovich transformation (valid only for $B_N > 0$):

$$\exp\left(\frac{B_N}{2N}\sum_{i,j}^N \phi_i\phi_j\right) = \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi/NB_N}} \exp\left(-\frac{NB_N y^2}{2} + B_N\sum_{i=1}^N \phi_i y\right). \tag{5.40}$$

This allows us to rewrite the probability distribution in Eq. (5.39) as

$$P(\boldsymbol{\varphi}_N) = \frac{1}{Z_N}\sqrt{\frac{NB_N}{2\pi}} \int_{-\infty}^{\infty} dy \prod_{i=1}^N \exp\left[-\frac{1}{2}\left(A_N + \frac{B_N}{N}\right)\phi_i^2 + (\mu + B_N y)\phi_i - \frac{NB_N}{2}y^2\right]. \tag{5.41}$$

The partition function then reads

$$Z_N = \left[\frac{2\pi}{\left(A_N + \frac{B_N}{N} \right)} \right]^{N/2} \sqrt{\frac{A_N N + B_N}{A_N N + B_N (N-1)}} \exp\left[\frac{N^2 \mu^2}{2(A_N N + B_N(N-1))} \right] \quad (5.42)$$

valid only for $B_N < A_N + B_N/N$. Eq. (5.42) can be used to derive the following relations:

$$\frac{\partial \log Z_N}{\partial \mu} = N\mathbb{E}[\phi_i|N], \quad (5.43)$$

$$\frac{\partial \log Z_N}{\partial A_N} = -\frac{N}{2}\mathbb{E}[\phi_i^2|N], \quad (5.44)$$

$$\frac{\partial \log Z_N}{\partial B_N} = \left(\frac{N-1}{2} \right) \mathbb{E}[\phi_i \phi_j|N] \quad \text{with} \quad (i \neq j) \quad (5.45)$$

which assuming symmetric volumes ($\mathbb{E}[\phi_i|N] = 0$, i.e. $\mu = 0$) are equivalent to

$$\mathbb{E}[\phi_i^2|N] = \frac{A_N + 2B_N/N - B_N}{(A_N + B_N/N - B_N)(A_N + B_N/N)} \quad (5.46)$$

and

$$\mathbb{E}[\phi_i \phi_j|N] = \frac{B_N/N}{(A_N + B_N/N - B_N)(A_N + B_N/N)}. \quad (5.47)$$

Furthermore, combining Eqs. (5.46) and (5.47) we can derive the volume correlation

$$\mathbb{C}_\phi(N) = \frac{\mathbb{E}[\phi_i \phi_j|N] - \mathbb{E}[\phi_i|N]^2}{\mathbb{E}[\phi_i^2|N] - \mathbb{E}[\phi_i|N]^2} = \frac{\mathbb{E}[\phi_i \phi_j|N]}{\mathbb{E}[\phi_i^2|N]} = \frac{B_N/N}{A_N + 2B_N/N - B_N}. \quad (5.48)$$

Vice versa, from Eqs. (5.46) and (5.48) we can obtain for the model parameters

$$A_N = \frac{1 - 2\mathbb{C}_\phi(N) + N\mathbb{C}_\phi(N)}{(1 - \mathbb{C}_\phi(N))(1 - \mathbb{C}_\phi(N) + N\mathbb{C}_\phi(N))\mathbb{E}[\phi_i^2|N]} \quad (5.49)$$

and

$$B_N = \frac{N\mathbb{C}_\phi(N)}{(1 - \mathbb{C}_\phi(N))(1 - \mathbb{C}_\phi(N) + N\mathbb{C}_\phi(N))\mathbb{E}[\phi_i^2|N]}, \quad (5.50)$$

which are useful to fit the Gaussian model to data. We will use Eq. (5.49) and Eq. (5.50) to estimate respectively A_N and B_N, replacing correlations and expectations by their empirical natural counterpart. The properties of this kind of estimators, belonging to the GMM (Generalized Method of Moments) is for instance discussed in Ref. [73].

5.A.3 Analytical computation of market impact

To compute analytically the market impact function $I_N(\phi)$ from Eq. (5.18) in the Gaussian correlated framework we adopt the following strategy

1. firstly we factorize the joint probability distribution $P(\boldsymbol{\varphi}_N)$ in Eq. (5.39) with a non-null external field $\mu \neq 0$;
2. secondly we use the trick of the previous point to compute the market impact $I_N(\phi)$ in presence of an effective field $\tilde{\mu}$ induced by the correlation of the net order flow $\phi_m = \sum_{i\neq k}^{N} \phi_i$ with the known metaorder of size ϕ.

Step 1. The joint probability distribution in Eq. (5.39) can be written in the following matrix form

$$P(\boldsymbol{\varphi}_N) = \frac{1}{Z_N} \exp\left(-\frac{1}{2}\boldsymbol{\varphi}_N^T \mathbb{M} \boldsymbol{\varphi}_N + \boldsymbol{\mu}^T \boldsymbol{\varphi}_N\right), \tag{5.51}$$

where

- $\mathbb{M}$ is a NxN real and symmetric matrix with the elements on the principal diagonal equal to A_N and the ones elsewhere equal to $-B_N/N$,
- $\boldsymbol{\mu}$ is a N-dimensional vector with all the elements equal to the scalar $\mu \neq 0$.

Through the orthogonal transformation $\tilde{\boldsymbol{\varphi}}_N = \mathbb{O}\boldsymbol{\varphi}_N$ which diagonalizes the matrix $\mathbb{M}$, i.e. $\mathbb{O}^T\mathbb{M}\mathbb{O} = \mathrm{diag}(\lambda_1, \cdots, \lambda_N)$, the joint probability distribution $P(\boldsymbol{\varphi}_N)$ factorizes as

$$P(\tilde{\boldsymbol{\varphi}}_N) = \frac{1}{Z_N} \prod_{m=2}^{N} \exp\left[-\frac{\lambda_2}{2}\tilde{\phi}_m^2\right] \exp\left[-\frac{1}{2}\lambda_1\tilde{\phi}_1^2 + \mu\sqrt{N}\tilde{\phi}_1\right] \tag{5.52}$$

where the N eigenvalues of the matrix $\mathbb{M}$

$$\lambda_1 = A_N - (N-1)\frac{B_N}{N} = \frac{1}{\mathbb{E}[\phi_i^2|N](1 - \mathbb{C}_\phi(N) + N\mathbb{C}_\phi(N))} \tag{5.53}$$

and

$$\lambda_2 = \lambda_3 = \cdots = \lambda_N = A_N + \frac{B_N}{N} = \frac{1}{\mathbb{E}[\phi_i^2|N](1 - \mathbb{C}_\phi(N))} \tag{5.54}$$

allow us to rewrite the partition function as

$$Z_N = \sqrt{\frac{2\pi}{\lambda_1}} \left[\sqrt{\frac{2\pi}{\lambda_2}}\right]^{N-1} \exp\left[\frac{N\mu^2}{2\lambda_1}\right]. \tag{5.55}$$

In particular, it emerges that the first component of $\tilde{\boldsymbol{\varphi}}_N = \mathbb{O}\boldsymbol{\varphi}_N$ is equal to

$$\tilde{\phi}_1 = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \phi_i, \tag{5.56}$$

which put in evidence that the orthogonal basis change $\boldsymbol{\varphi}_N \to \tilde{\boldsymbol{\varphi}}_N$ is a useful trick to compute the market impact in the context of correlated Gaussian metaorders.

Step 2. To calculate the market impact $I_N(\phi)$ defined in Eq. (5.18) with N overall correlated Gaussian metaorders and in absence of an external field it is necessary to explicit the conditional probability distribution

$$P(\boldsymbol{\varphi}_N|\phi_k=\phi)=\frac{P(\phi_1,\cdots,\phi,\cdots,\phi_N)}{p(\phi)} \tag{5.57}$$

where $P(\phi_1,\cdots,\phi,\cdots,\phi_N)$ is given by Eq. (5.39) setting $\mu=0$ while the marginal one is equal to

$$p(\phi)=\frac{1}{\sqrt{2\pi\tilde{\lambda}_1/(\lambda_1\lambda_2)}}\exp\left[-\frac{\phi^2}{2}\frac{\lambda_1\lambda_2}{\tilde{\lambda}_1}\right], \tag{5.58}$$

with λ_1 and λ_2 respectively given by Eqs. (5.53) and (5.54) and

$$\tilde{\lambda}_1=A_N-\frac{N-2}{N}B_N=\frac{1}{\mathbb{E}[\phi_i^2|N][1-\mathbb{C}_\phi(N)+N\mathbb{C}_\phi(N)][1-\mathbb{C}_\phi(N)]}\,. \tag{5.59}$$

It follows that the conditional probability distribution is equal to

$$P(\boldsymbol{\varphi}_N|\phi_k=\phi)=\underbrace{\frac{\exp\left[-\frac{(N-1)B_N^2\phi^2}{2N^2\tilde{\lambda}_1}\right]}{(2\pi/\tilde{\lambda}_1)^{1/2}(2\pi/\lambda_2)^{\frac{N-2}{2}}}}_{\Theta_N^{-1}(\phi)}\times\exp\left[-\frac{A_N}{2}\sum_{i\neq k}^N\phi_i^2+\frac{B_N}{N}\sum_{\substack{i<j\\ i,j\neq k}}^N\phi_i\phi_j+\underbrace{\frac{\phi B_N}{N}}_{\tilde{\mu}}\sum_{i\neq k}^N\phi_i\right] \tag{5.60}$$

where it emerges that the conditioning to the metaorder with volume $\phi_k=\phi$ is equivalent to the introduction of an effective field $\tilde{\mu}$ proportional to ϕ. This implies that the price impact in Eq. (5.18) is given solving the following conditional expectation

$$\begin{aligned}
I_N(\phi) &= \mathbb{E}[\mathcal{I}(\boldsymbol{\varphi}_N)|\phi_k=\phi]= \\
&= \int_{-\infty}^{\infty}\prod_{i\neq k}^N \mathrm{d}\phi_i\frac{P(\phi_1,\cdots,\phi,\cdots,\phi_N)}{p(\phi)}\left(\phi+\sum_{i\neq k}^N\phi_i\right)^{\bullet 1/2}= \\
&= \frac{1}{\Theta_N(\phi)}\int_{-\infty}^{\infty}\prod_{i\neq k}^N \mathrm{d}\phi_i\exp\left[-\frac{A_N}{2}\sum_{i\neq k}^N\phi_i^2+\frac{B_N}{N}\sum_{\substack{i<j\\ i,j\neq k}}^N\phi_i\phi_j+\tilde{\mu}\sum_{i\neq k}^N\phi_i\right]\times \\
&\times\left(\phi+\sum_{i\neq k}^N\phi_i\right)^{\bullet 1/2}= \\
&= \frac{1}{\Theta_N(\phi)}\int_{-\infty}^{\infty}\prod_{i\neq k}^N \mathrm{d}\phi_i\exp\left[-\boldsymbol{\varphi}^{*T}\mathbb{M}^*\boldsymbol{\varphi}^*+\tilde{\mu}\boldsymbol{\varphi}^*\right]\left(\phi+\sum_{i\neq k}^N\phi_i\right)^{\bullet 1/2}\,.
\end{aligned} \tag{5.61}$$

Herein

- $\boldsymbol{\varphi}^* = \{\phi_i\}_{i=1,\cdots,N}^{i\neq k}$ is a vector that contains the $N-1$ unknown metaorders volumes simultaneously executed with the one known $\phi_k = \phi$,

- $\mathbb{M}^*$ is a $(N-1)\times(N-1)$ symmetric and real matrix with A_N on the principal diagonal and $-B_N/N$ elsewhere: it is easy to check that its eigenvalues are respectively $\lambda_1^* = \tilde{\lambda}_1$ as in Eq. (5.59) and $\lambda_m^* = \lambda_2$ as in Eq. (5.54) for $m = 2, \cdots, N-1$.

As mentioned before, to solve Eq. (5.61) it is useful to use the trick described in Step 1: we apply in Eq. (5.61) the orthogonal transformation $\tilde{\boldsymbol{\phi}}^* = \mathbb{H}\boldsymbol{\varphi}^*$ that diagonalizes the matrix $\mathbb{M}^*$ (i.e. $\mathbb{H}^T\mathbb{M}^*\mathbb{H} = \mathrm{diag}(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_{N-1})$) and since the determinant of the Jacobian matrix associated to this transformation is equal to one, we obtain that

$$I_N(\phi) = \frac{1}{\Theta_N(\phi)} \int_{-\infty}^{\infty} \prod_{m=2}^{N} d\tilde{\phi}_m^* \exp\left[-\frac{\lambda_2}{2}(\tilde{\phi}_m^*)^2\right] \times$$
$$\times \int_{-\infty}^{\infty} d\tilde{\phi}_1^* \exp\left[-\frac{1}{2}\tilde{\lambda}_1(\tilde{\phi}_1^*)^2 + \tilde{\mu}\sqrt{N-1}\tilde{\phi}_1^*\right](\phi + \sqrt{N-1}\tilde{\phi}_1^*)^{\bullet 1/2}; \quad (5.62)$$

then integrating in $\tilde{\phi}_m^*$ for $m = 2, \cdots, N-1$, completing the square in the argument of the $\exp[-\frac{1}{2}\tilde{\lambda}_1(\tilde{\phi}_1^*)^2 + \tilde{\mu}\sqrt{N-1}\tilde{\phi}_1^*]$ and doing the variable change $x = \phi + \sqrt{N-1}\tilde{\phi}_1^*$ we derive the following final expression

$$I_N(\phi) = \frac{1}{\sqrt{2\pi(N-1)/\tilde{\lambda}_1}} \int_0^{\infty} \mathrm{d}x\sqrt{x} \left(\exp\left[-\frac{\tilde{\lambda}_1}{2(N-1)}\left(x - \phi(1+(N-1)\mathbb{C}_\phi(N))\right)^2\right] + \right.$$
$$\left. - \exp\left[-\frac{\tilde{\lambda}_1}{2(N-1)}\left(x + \phi(1+(N-1)\mathbb{C}_\phi(N))\right)^2\right]\right). \quad (5.63)$$

Remark 4. From Eq. (5.63) it follows that the price impact $I_N(\phi)$ in the correlated Gaussian framework is equivalent to the one computed in the independent Gaussian case (see Eqs. (5.21) and (5.22)) substituting

$$\phi \longrightarrow \phi[1 + (N-1)\mathbb{C}_\phi(N)] \quad (5.64)$$

and

$$\Sigma_N^2 \longrightarrow 1/\tilde{\lambda}_1. \quad (5.65)$$

5.A.4 Market impact with correlated signs and i.i.d. unsigned volumes

In Section 5.5.4 we have presented a general model in which the metaorder signs $\epsilon_i = \pm 1$ are correlated while the unsigned volumes $|\phi_i|$ are i.i.d. and described by a generic distribution $p(|\phi_i|)$ defined on the finite positive support $(0, 1)$. This means that the joint probability distribution is factorizable as

$$P(\boldsymbol{\varphi}_N) = \mathcal{P}(\boldsymbol{\epsilon}_N)\prod_{i=1}^{N} p(|\phi_i|). \quad (5.66)$$

In this theoretical setup we are not able to compute analytically the market impact $I_N(\phi)$ and we will use numerical simulations. To this aim we introduce a latent discrete variable $\tilde{\epsilon}$ in order to simulate N correlated signs with the following statistical model

$$\mathbb{P}(\epsilon_i|\tilde{\epsilon}) = \frac{1}{2}(1+\gamma_\epsilon \epsilon_i \tilde{\epsilon}) \tag{5.67}$$

where γ_ϵ can be estimated from data by averaging the realized sign correlation ρ_ϵ appearing in Eq. (5.16), i.e.

$$\mathbb{E}[\rho_\epsilon|N] = \gamma_\epsilon^2. \tag{5.68}$$

For clarity we explicitly omit the γ_ϵ's dependence on N. Thus to simulate the model we fix $\epsilon_k = +1$, we draw a hidden factor $\tilde{\epsilon}$ from

$$\mathbb{P}(\tilde{\epsilon}|\epsilon_k = +1) = \frac{1}{2}(1+\gamma_\epsilon \tilde{\epsilon}) \tag{5.69}$$

and then we sample $N-1$ other correlated signs $\{\epsilon_i\}_{i=1,\cdots,N}^{i\neq k}$ with probability

$$\mathbb{P}(\epsilon_i|\tilde{\epsilon}) = \frac{1}{2}(1+\gamma_\epsilon \epsilon_i \tilde{\epsilon})\,. \tag{5.70}$$

Numerical computation of market impact

We summarize the main steps for the numerical calibration of the market impact $I_N(\phi)$ from data. Given the number N of metaorders per stock/day pair and fixing $|\phi_k|= \phi > 0$:

1. We compute the average sign correlation $\mathbb{C}_\epsilon(N) = \mathbb{E}[\rho_\epsilon|N]$ as to obtain γ_ϵ through Eq. (5.68).
2. After fixing the direction $\epsilon_k = +1$ we simulate $N-1$ correlated signs using Eq. (5.70).
3. We sample $N-1$ random variables $|\phi_i|$ from an half-normal distribution with mean $\Sigma_N\sqrt{2/\pi}$ and standard deviation $\Sigma_N\sqrt{1-2/\pi}$ where Σ_N represents the empirical standard deviation of signed volumes.
4. We compute numerically the market impact $I_N(\phi) = \mathbb{E}[\mathcal{I}(\boldsymbol{\varphi}_N)|\epsilon_k = +1, |\phi_k|= \phi]$ where $\mathcal{I}(\boldsymbol{\varphi}_N) = \left(\sum_{i=1}^N \phi_i\right)^{\bullet 1/2}$ and $\phi_i = \epsilon_i|\phi_i|$, as defined in Eq. (5.6).
5. Finally, we compute $I(\phi)$ averaging $I_N(\phi)$ over the empirical distribution $p(N)$ shown in the left panel of Figure 5.1.

Chapter 6

Slow decay of impact in equity markets

6.1 Introduction

It is now well documented that a series of buy (sell) trades, originating from the same institution, pushes on average the price upwards (downwards), by a quantity proportional to the square root of the total volume Q of the buy trades – see e.g. [2, 5, 9, 10, 12, 13, 37, 38, 50, 82]. What happens when such a series of trades belonging to the same trading decision is completed? One expects that once the extra buy (sell) pressure subsides, impact will revert somewhat. There is however no consensus on the detailed behaviour of such impact decay. The long-run asymptotic price after the metaorder is expected to depend on the information on which trading is based, so it is customary to decompose the observed impact into a reaction impact, that is a mechanical property of markets, unrelated to information, and a prediction impact that depends on the quality of information contained in the trade [34].

From a modeling point of view, several hypotheses have been put forward. In the stylized, fair pricing theory of Farmer, Gerig, Lillo & Waelbroeck (FGLW) [15], an equilibrium condition is derived between liquidity providers and a broker aggregating informed metaorders from several funds, in which the average price payed during the execution is equal to the price at the end of the reversion phase. If metaorder size distribution is power law with tail exponent 3/2, the observed impact is predicted to decay towards a plateau value whose height is 2/3 of the peak impact, i.e. the impact reached exactly when the metaorder execution is completed. Within the latent order book model [2, 7], reaction impact is expected to decay as a power-law of time, reaching a (small) asymptotic value after times corresponding to the memory time of the market [69]. A similar behaviour is predicted by the propagator model [18]. Note that while the latent order book model predicts that the permanent reaction impact is linear in Q [69] (in agreement with no-arbitrage arguments [25, 87]), the FGLW theory implies that permanent impact scales as the peak impact, i.e. as $\sqrt{Q}$.

As far as empirical data is concerned, the situation is rather confusing, mostly because the determination of the time when the relaxation terminates is not unique. Some papers, determining permanent impact shortly after the end of the metaorder, have reported results compatible with the FGLW 2/3 factor [5, 8, 10, 11, 88], although Gomes & Waelbroeck [11] note that the impact of uninformed trades appears to relax to zero. Brokmann et al. [12], on the other hand, underline the importance of metaorders split over many successive days, as this may strongly bias upwards the apparent plateau value. After accounting for both metaorder correlations and prediction impact, Brokmann et al. [12] conclude that impact decays as a power-law over several days, with no clear asymptotic value. The work of Bacry et al. [9] leads to qualitatively similar conclusions.

In the present chapter, we revisit this issue using the ANcerno database with a closer focus on impact decay. Extending the time horizon beyond that considered in [10], we establish unambiguously that impact decays below the 2/3 plateau, which is observed as average value of the impact at the end of the same day of the metaorder execution. Specifically, we find that the overnight contribution to impact decay is small, in agreement with the idea that the decay takes place in volume time rather than in calendar time. After accounting for metaorder temporal correlations, impact decay is well fitted by a power-law for intraday time scales and an exponentially truncated power-law for multiday horizons, extrapolating to a plateau value $\approx 1/3$ of the peak impact beyond several weeks. For such long time lags, however, market noise becomes dominant and makes it difficult to conclude on the asymptotic value of impact, which is a proxy for the (long time) information content of the trades in the ANcerno database.

6.2 Data and definitions

We use the heterogeneous dataset provided by ANcerno (for a details see Chapter 3). Its structure allows the identification of metaorders relative to the trading activity of a diversified pool of anonymized institutional investors. It follows that each metaorder is characterized by a stock symbol, the total volume Q (in number of shares) and the times at the start t_s and at the end t_e of its execution with sign $\epsilon = \pm 1$ (buy/sell). Our sample covers for a total of 880 trading days, from January 2007 to June 2010 and we select only stocks in the Russell 3000 index. The cleaning procedure introduced in Ref. [10] is applied to remove possible spurious data. In this way the available sample is represented by around 8 million metaorders distributed quite uniformly in time and across market capitalizations representing around 5% of the total daily market volume.

Let us briefly recall the main observables useful to describe the metaorder execution and the price relaxation, namely the participation rate η and the duration D measured in volume time. The participation rate η is defined as the ratio between the number of shares Q traded by the metaorder and the whole market volume traded during the execution time interval $[t_\mathrm{s}, t_\mathrm{e}]$

$$\eta = \frac{Q}{V(t_\mathrm{e}) - V(t_\mathrm{s})}, \tag{6.1}$$

where $V(t)$ is the cumulative volume transacted in the market between the start of the trading day and time t. The metaorder duration in volume time D is expressed as

$$D = \frac{V(t_{\mathrm{e}}) - V(t_{\mathrm{s}})}{V_{\mathrm{d}}}, \tag{6.2}$$

where V_{d} is the total daily market volume. The unsigned daily fraction $\phi = Q/V_{\mathrm{d}}$ is then given by

$$\phi = \eta \times D. \tag{6.3}$$

For a description of the metaorder statistics the interested reader can find details in the Chapter 3.

To investigate the price relaxation process for $t > t_{\mathrm{e}}$ we introduce the following main definitions of market impact. Let us suppose that an asset manager decides to buy or to sell a metaorder of Q shares sending it at time $t = t_{\mathrm{s}}$ to a broker or to an execution algorithm where it is executed sequentially in smaller orders on market until to completion at time $t = t_{\mathrm{e}}$. The market impact is then defined in terms of the rescaled log-price $s(t) := (\log S(t))/\sigma_{\mathrm{d}}$, where $S(t)$ is the average market mid-price at time t and $\sigma_{\mathrm{d}} = (S_{\mathrm{high}} - S_{\mathrm{low}})/S_{\mathrm{open}}$ is a noisy estimator of the daily volatility, estimated from the daily high, low and open prices. Given the rescaled average market mid-price at the start of the metaorder $s_{\mathrm{start}} = s(t_{\mathrm{s}})$ and at the end of its execution, $s_{\mathrm{end}} = s(t_{\mathrm{e}})$, we quantify its Start-to-End price impact I_{SE} with the following antisymmetric expectation

$$I_{SE}(\phi) = \mathbb{E}[\epsilon \cdot (s_{\mathrm{end}} - s_{\mathrm{start}})|\phi] \tag{6.4}$$

where $\epsilon = \pm 1$ is the direction of the metaorder (buy/sell). In practice, we compute the market impact curve $I_{SE}(\phi)$ by dividing the data into evenly populated bins according to the volume fraction ϕ and computing the conditional expectation of impact for each bin [9, 10, 41]. Henceforth, error bars are determined as standard errors. Similarly, we will define the Start-to-Close impact I_{SC} by replacing in Eq. (6.5) the end price $s(t_{\mathrm{e}})$ by the close log-price of the day $s_{\mathrm{close}} = s(t_{\mathrm{c}})$, i.e.

$$I_{SC}(\phi) = \mathbb{E}[\epsilon \cdot (s_{\mathrm{close}} - s_{\mathrm{start}})|\phi]. \tag{6.5}$$

6.3 Intraday impact and post-trade reversion

In Figure 6.1, we show the Start-to-End impact I_{SE} and Start-to-Close impact I_{SC} as a function of the daily volume fraction ϕ. Clearly, I_{SE} behaves as a square root of the volume fraction ϕ in an intermediate regime $10^{-3} \lesssim \phi \lesssim 10^{-1}$, as reported in many previous studies [2, 5, 9, 10, 12, 13, 37, 38, 50, 82]. For smaller volume fractions, impact is closer to linear [10] – see Chapter 4 for a discussion of this effect. The Start-to-Close impact, measured using exactly the same metaorders, is below the Start-to-End impact ($I_{SC} < I_{SE}$), showing that some post-trade reversion has taken place between the metaorder completion time t_{e} and the market close time t_{c}.

The ratio between these two impact curves is plotted in Figure 6.1 (right panel). Interestingly, the mean value over all ϕ is found to be 0.66 ± 0.04, in close agreement

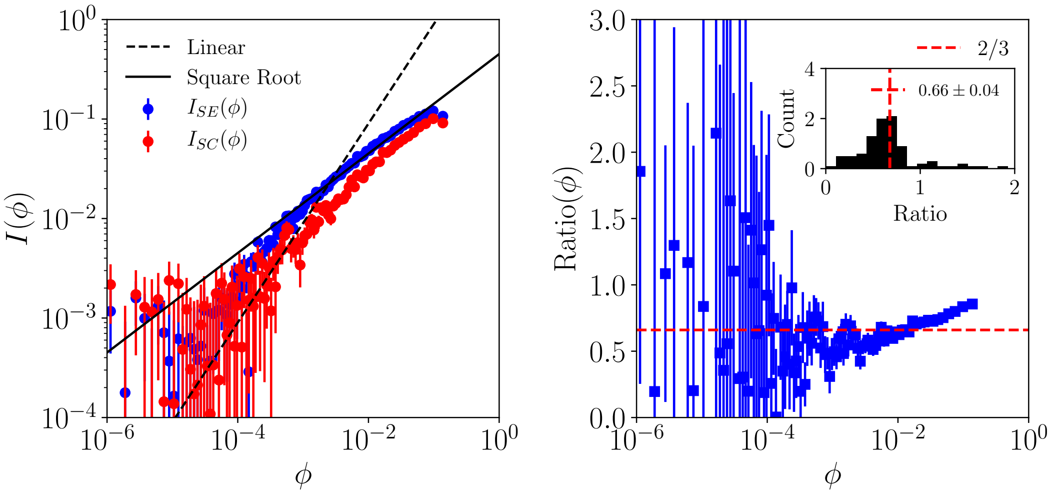

Figure 6.1: (Left panel) Start-to-End impact I_{SE} and Start-to-Close impact I_{SC} as a function of the daily volume fraction ϕ. We also show the square-root impact law $I(\phi) \propto \sqrt{\phi}$ (plain line) and a linear impact law (dotted line). The slope of I_{SC} appears to be larger than that of I_{SE} as a consequence of a stronger impact decay contribution for smaller ϕ's. (Right panel) The ratio $I_{SC}(\phi)/I_{SE}(\phi)$, computed in each volume fraction bin ϕ. Its average over all ϕ is $= 0.66 \pm 0.04$. The empirical distribution of the ratio is presented in the inset. Note that for $\phi \gtrsim 10^{-3}$, this ratio increases with ϕ. In both the panels the error bars are standard errors.

with the 2/3 ratio predicted by FGLW, thus confirming previous empirical findings [5, 8, 10, 11, 88]. However, a closer look at the plot reveals that the ratio systematically increases as ϕ increases. Since larger metaorders (i.e. large ϕ) tend to take longer to execute, one expects the End-to-Close time $T_{EC} = t_c - t_e$ to decrease as ϕ increases. Therefore impact decay between the end of the metaorder and the end of the day should be, on average, larger for small order size ϕ.

In order to validate this hypothesis, we now characterize the intraday price reversion by computing the ratio $R(z) = I_{SC}/I_{SE}$ as a function of the variable $z = V_{EC}/V_{SE}$, where $V_{EC} = V(t_c) - V(t_e)$ and $V_{SE} = V(t_e) - V(t_s)$ are respectively the total market volume executed in the time intervals T_{EC} and T_{SE} (similar results – not shown – are obtained as a function of $z' = T_{EC}/T$, where $T = t_e - t_s$ is the metaorder execution time). The results are shown in Figure 6.2 (left panel). One clearly sees that impact decays continuously as z increases, and is in fact well fitted by the prediction of the propagator model [18, 34], namely $\mathcal{I}_{\text{prop}}(z) = (1+z)^{1-\beta} - z^{1-\beta}$ with $\beta = 0.22$ [1]. Interestingly, if one restricts to a smaller interval $z \in [0, 2]$, as in [10, 88], one finds that the decay appears to saturate around the 2/3 value (see inset in Figure 6.2), but zooming out leaves no

[1]To note, β is the decay exponent of the propagator, $G(t) \sim t^{-\beta}$, see e.g. [34].

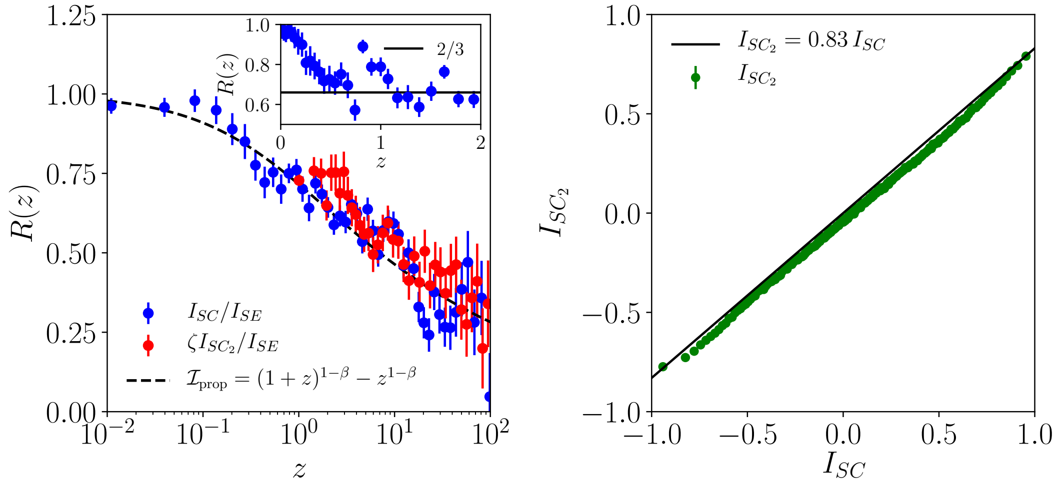

Figure 6.2: (Left panel) Price relaxation $R(z)$ over two consecutive days. Blue points: impact decay within the same day of the metaorder's execution, i.e. $R(z) = I_{SC}/I_{SE}$ as a function of $z = V_{EC}/V_{SE}$, in a semi-log scale. Red points: impact decay using the close of the next day, i.e. $R(z) = \zeta\, I_{SC_2}/I_{SE}$ as a function of $z = V_{EC_2}/V_{SE}$, with $\zeta = 0.80$. Both sets of points are well fitted by the prediction of the propagator model: $\mathcal{I}_{\text{prop}}(z) = (1+z)^{1-\beta} - z^{1-\beta}$ with $\beta = 0.22$. Inset: same day impact decay, in a linear-linear plot restricted to $z \in [0, 2]$, suggesting relaxation towards a $2/3$ value (horizontal line). The error bars are standard errors. (Right panel) Average of Start-to-next day Close I_{SC_2}, conditioned to different values of I_{SC}. The regression lines yield $I_{SC_2} = 0.83\, I_{SC}$.

doubt that impact is in fact decaying to smaller values. In Appendix 6.A more insights on the intraday price relaxation are discussed.

6.4 Next day reversion

Quite interestingly, impact decays much in the same way over the next day: in the same figure (left panel in Figure 6.2) we plot $R(z) \simeq I_{SC_2}/I_{SE}$ as a function of $z = V_{EC_2}/V_{SE}$, where C_2 refers to the close of the next day and $V_{EC_2} = V_{EC} + V_{\text{d}}$ (i.e. the overnight does not contribute to T_{EC_2}). Provided one applies a factor $\zeta \approx 4/5$ that accounts for the autocorrelation of metaorders (see next section, and left panel in Figure 6.3)[2], the next day impact decay nicely falls in the continuation of the intraday decay, and is also well accounted for by the very same scaling function $\mathcal{I}_{\text{prop}}(z)$.

Figure 6.2 (right panel) provides complementary information: we show the average

[2]More precisely, we have set $\zeta = 1/(1 + C(1))$, where $C(\tau)$ is the autocorrelation function plotted in Figure 6.3 (left panel).

of I_{SC_2} conditioned to different values of I_{SC}, which clearly demonstrates that these two quantities are proportional and related to the same decay mechanism. It implies in particular that the impact measured at the end of the next day would still behave as a square root of the volume of the executed metaorder.

6.5 Impact decay over multiple days

Having established that impact decay occurs both intraday and during the next day, it is tempting to conjecture that impact will continue to decay on longer time scales. However, the empirical investigation of such a decay faces several hurdles. First, as the time lag increases, the amount of noise induced by overall market moves becomes larger and larger (in fact as $\sigma_{\mathrm{d}}\sqrt{\tau}$, where τ is the number of days). Second, metaorders are often executed over several days, leading to long range autocorrelations of the order flow. This effect, investigated considering metaorders from the same fund [34], is here investigated by considering a very heterogeneous set of funds and illustrated in the left panel of Figure 6.3. We find that metaorder signs autocorrelation is well fitted by an exponentially truncated power law with a time scale of ≈ 26 days and an exponent γ fixed by the propagator model constraint $\gamma = 1 - 2\beta$. Intuitively, these correlations may mask the decay of impact, as trades in the same direction during the following days tend to counterbalance impact reversion, leading to an apparent increase of impact (see Figure 6.3, right panel, inset). This contribution should be somehow removed to estimate the natural decay of impact.

A possible way to overcome the latter problem is to apply a deconvolution method introduced by Brokmann et al. [12]. We will study our metaorder database at the daily time scale with the same angle: for each day $\tau = \{1, \cdots, 880\}$ and asset we computed the net daily traded volume $\Phi(\tau) = \sum_{i=1}^{N} \epsilon_i \phi_i$ where N is the number of metaorders in the database, for a given asset and a given day τ. As discussed in Chapter 5, the impact of a set of different metaorders, all executed the same day and on the same asset, is well described by an extended square-root law where all metaorders are bundled together:

$$\mathcal{I}(\Phi) = Y \times \sigma_{\mathrm{d}} \times \Phi^{\bullet 1/2}, \tag{6.6}$$

where we use the signed power notation $x^{\bullet 1/2} := \mathrm{sign}(x) \times \sqrt{|x|}$.

The return of the asset between the last day close and the close of each day τ is denoted $r(\tau)$. The method used in Ref. [12] amounts to assuming a quasi-linear model, i.e.

$$r(\tau) = \beta_{\mathrm{capm}}(\tau)\, r_{\mathrm{M}}(\tau) + \sigma_{\mathrm{d}} \sum_{\ell=0}^{L} G(\ell)\, \widetilde{\Phi}^{\bullet 1/2}(\tau - \ell) + \xi(\tau), \tag{6.7}$$

where $G(\ell)$ are coefficients, L is a certain horizon (taken to be $L = 50$ days), ξ is a noise term and $\beta_{\mathrm{capm}}(\tau)\, r_{\mathrm{M}}(\tau)$ is the systematic component that takes into account the market drift: $\beta_{\mathrm{capm}}(\tau)$ is the beta of the traded stock computed on the period from $\tau - 20$ to $\tau + 20$ and $r_{\mathrm{M}}(\tau)$ the daily close-close return of the market (here the Russell

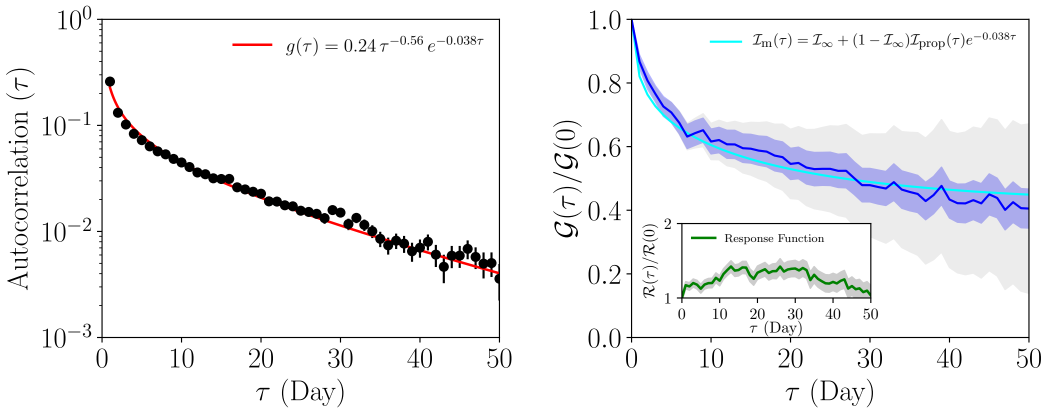

Figure 6.3: (Left panel) Empirical autocorrelation of the signed square-root volume imbalance $\Phi^{\bullet 1/2}$, as a function of the lag τ, averaged over all stocks. This autocorrelation persists over many days, as it is fitted as an exponentially truncated power law $g(\tau) = a\tau^{-\gamma}e^{-b\tau}$ with $a = 0.24 \pm 0.04$ and $b = 0.038 \pm 0.002$ (corresponding to $1/b \simeq 26$ days). The value of the exponent γ is fixed to $1 - 2\beta = 0.56$, as dictated by the propagator model [18]. (Right panel) Normalized impact kernel $\mathcal{G}(\tau)/\mathcal{G}(0)$ estimated using Eq. (6.8) for $\tau \in [1, 50]$ days. The fit corresponds to the exponentially truncated modified propagator model $\mathcal{I}_{\rm m}(\tau)$ with $b = 0.038$ (see left panel), which provides an asymptotic decay level $\mathcal{I}_\infty \approx 0.42 \pm 0.01$. The error bars on the graph are (i) bootstrap errors (blue region) and (ii) cumulated regression errors (grey region). Inset: normalized response function $\mathcal{R}(\tau)/\mathcal{R}(0)$ as a function of τ (see definition in Eq. (6.10)).

3000 index). Finally, $\widetilde{\Phi}^{\bullet 1/2}(\tau) = \Phi^{\bullet 1/2}(\tau) - \beta_{\rm capm}(\tau)\langle\Phi^{\bullet 1/2}(\tau)\rangle_{\rm stocks}$, where we subtract $\beta_{\rm capm}$ times the cross-sectional average of the expected impact[3].

Pooling all the stocks together[4], a least-square regression allows us to determine the coefficients $G(\ell)$, from which we reconstruct the reactional impact kernel $\mathcal{G}(\tau)$ as

$$\mathcal{G}(\tau) = \sum_{\ell=0}^{\tau} G(\ell). \tag{6.8}$$

The kernel $\mathcal{G}(\tau)$ is a proxy of the impact of an isolated metaorder. If the metaorder was uniformed, $\mathcal{G}(\tau)$ would describe the mechanical reaction of the market to such a trade. Any non-zero asymptotic value of $\mathcal{G}(\tau \to \infty)$ would either reveal that metaorders are on average informed, or that even random trades have positive permanent impact on prices (as in Ref. [69]).

[3] Note that in the quasi-linear model described by Eq. (6.7) we use $\tilde{\Phi}^\bullet$ instead of $\Phi^\bullet$ in such a way to remove the cross-sectional market impact contribution given by $\beta_{\rm capm}(\tau)\langle\Phi^{\bullet 1/2}(\tau)\rangle_{\rm stocks}$: for consistency we introduce the intercept $\beta_{\rm capm}(\tau)\,r_{\rm M}(\tau)$ on the right side of Eq. (6.7).

[4] We have checked that different subsamples of the full sample lead to similar results (for example, slicing the pool of stocks according to their market capitalisation, see Figure 6.4).

To estimate error bars, we generated 200 bootstrap samples using all 1500 stocks, and ran the linear regression Eq. (6.7) on each of them. The average result is shown in Figure 6.3 (right panel), together with error bars coming from the least-square regression and from the bootstrap procedure. From this graph, we see that the estimated impact kernel $\mathcal{G}(\tau)$ slowly decays in a time window comparable with the one over which we measure a persistent autocorrelation (as shown in the left panel of Figure 6.3). We have fitted the empirically determined, normalized impact kernel $\mathcal{I}_{\mathrm{m}}(\tau) := \mathcal{G}(\tau)/\mathcal{G}(0)$ using an ad-hoc modified propagator kernel, that accounts for a final exponential decay towards an asymptotic value $\mathcal{I}_\infty$:

$$\mathcal{I}_{\mathrm{m}}(\tau) := \mathcal{I}_\infty + (1 - \mathcal{I}_\infty)\mathcal{I}_{\mathrm{prop}}(\tau)e^{-b\tau}, \tag{6.9}$$

where b is a parameter fixed by the corresponding decay of the flow autocorrelation, see Figure 6.3 (left panel). Keeping the same shape for $\mathcal{I}_{\mathrm{prop}}(\tau)$ as the one describing the short-term decay of impact (i.e. fixing $\beta = 0.22$), the one-parameter fit gives $\mathcal{I}_\infty \approx 0.42$. Leaving b free in a 2-parameter fit leads to very similar values: $b = 0.03 \pm 0.01$ and $\mathcal{I}_\infty = 0.39 \pm 0.05$. However, setting $b = 0$ and leaving β free leads to $\beta = 0.15 \pm 0.04$ and a zero asymptotic value $\mathcal{I}_\infty = 0.0 \pm 0.19$.

Although the error bars are already large for $\tau = 50$, the fit seems to favor a non-zero asymptotic value $\mathcal{I}_\infty \approx 1/2$. Since the impact has on average already decayed to approximately 2/3 of its peak value at the end of the trading day, this value of $\mathcal{I}_\infty$ suggests a long time asymptotic plateau at $2/3 \times 1/2 \approx 1/3$ of the peak value, significantly below the 2/3 value predicted by FGLW (see also Figure 6.2). This can be taken as a measure of the information content of the trades in the ANcerno database. Since we have no knowledge about the intensity of the trading signal which triggered the metaorders, we cannot subtract the alpha component from the observed returns, as was done in Ref. [12], where after removing the alpha of the manager and the contribution of correlated trades, impact was found to decay to ≈ 0.15 of its initial value after 15 days. Adding to the regressors of Eq. (6.7) past values of $(r - \beta_{\mathrm{capm}} r_{\mathrm{M}})$, as a proxy for mean-reversion and/or trending signals that investors commonly use, we find a slightly larger plateau (≈ 0.54) when b is kept at the value 0.038. This reveals how noisy the data is, because one would have expected a decrease of $\mathcal{I}_\infty$ when including more alpha signal in the regression. However, we do observe mean reversion on short time scales and momentum beyond, as expected.

Finally, we also show in Figure 6.3 (right panel, inset) the full response function $\mathcal{R}(\tau)$, defined as [18]:

$$\mathcal{R}(\tau) := \Big\langle \sum_{\tau'=0}^{\tau} \widetilde{r}(\ell + \tau')\widetilde{\Phi}^{\bullet 1/2}(\ell) \Big\rangle_\ell, \tag{6.10}$$

where $\widetilde{r}(\tau) := r(\tau) - \beta_{\mathrm{capm}}(\tau)\, r_{\mathrm{M}}(\tau)$ and the average operation $\langle \cdots \rangle_\ell$ is done over all the days ℓ. This quantity elicits an *apparent* evolution of impact, without accounting for metaorder autocorrelations. Such autocorrelations are strong enough to make $\mathcal{R}(\tau)$ increase as a function of τ (see [18, 34] for comparable results). This plot illustrates how the autocorrelation of order flow can strongly bias the estimation of impact decay and its asymptotic value (see [12] for a similar discussion).

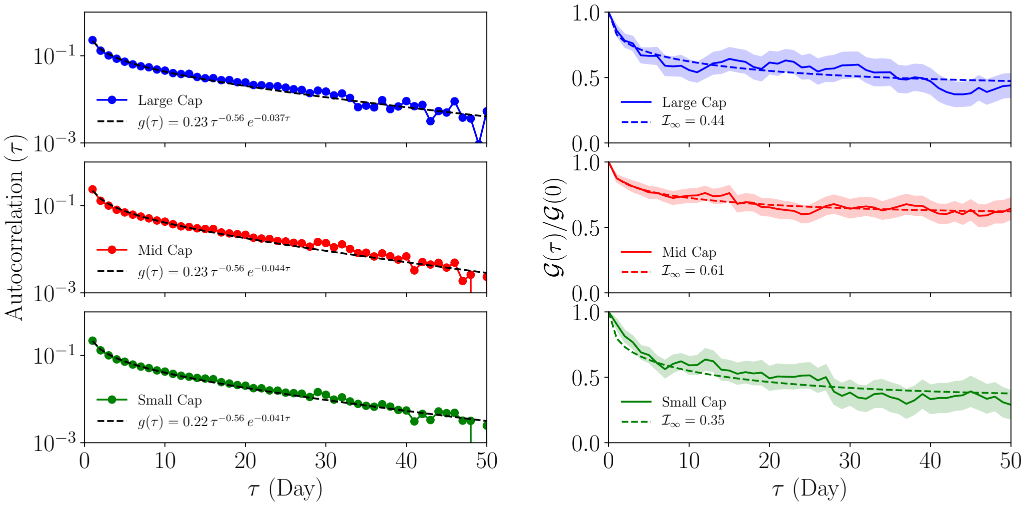

Figure 6.4: (Left panel) Empirical autocorrelation of the signed square root volume imbalance $\Phi^{\bullet 1/2}$, as a function of the lag τ, averaged over all stocks in a given market capitalisation tranche. Each function is fitted as an exponentially truncated power law $g(\tau) = a\tau^{-\gamma}e^{-b\tau}$ with $\gamma = 1 - 2\beta = 0.56$ (fixing $\beta = 0.22$). The parameters a and b are very close in the three cases. (Right panel) Normalized decay kernel $\mathcal{G}(\tau)/\mathcal{G}(0)$ estimated using Eq. (6.8) for $\tau \in [1, 50]$ days, again using all stocks in a given market cap tranche. The fit corresponds to the exponentially truncated modified propagator model $\mathcal{I}_\mathrm{m}(\tau)$ with $b = 0.038$, which provides an asymptotic decay level $\mathcal{I}_\infty \approx 0.44$ (large cap), $\mathcal{I}_\infty \approx 0.61$ (mid cap) and $\mathcal{I}_\infty \approx 0.35$ (small cap), all within the grey region of right panel in Figure 6.3. The error bars in the right panels are bootstrap errors.

6.6 Conclusions

In this chapter we presented an empirical study of the impact relaxation of metaorders executed by institutional investors in the U.S. equity market. We have shown that relaxation takes place as soon as the metaorder ends, and continues the following day with no apparent saturation at the plateau value corresponding to the fair pricing FGLW theory [15]. For example, the impact measured at the next-day close is, on average, around 4/5 of the impact at the end of the day when the metaorder is executed. The decay of impact is described by a power-law function at short time scales, while it appears to converge to a non-zero asymptotic value at long time scales ($\sim$ 50 days), equal to 1/2 of the impact at the end of the first day, which is $\approx$ 1/3 of the peak impact. Due to a significant, multiday correlation of the sign of executed metaorders, a careful deconvolution of the observed impact must be performed to extract the reaction impact contribution (where, possibly, some information contribution remains). Once this is done, our results match qualitatively those of Ref. [12], obtained on a smaller set of metaorders executed by a single manager (CFM). In particular, we find no support for the prediction of Farmer et al. [15], that the permanent impact equal to 2/3 of the peak impact.

Executing a quantity Q moves the price, on average, as $I(Q) = Y\sigma_{\rm d}\sqrt{Q/V_{\rm d}}$, where Y is a certain numerical constant [2,5,9,10,12,13,37,38,50,82]. Assuming that this impact is fully transient and decays back to zero at long times, the corresponding average cost of trading is $2/3\, I(Q)$. If the investor predicts a certain price variation Δ, his/her optimal trade size is given by the following maximization problem:

$$Q^* = \text{argmax}\left[\Delta Q - \frac{2}{3} Q\, I(Q)\right] \quad \Rightarrow \quad I(Q^*) = \Delta. \tag{6.11}$$

The last equation means that the investor should trade until his/her average impact pushes the price up to the predicted level Δ, but not beyond. For truly informed investors, there should be no decay of impact at all, since the price has been pushed to its correctly predicted value. For uninformed investors, on the other hand, impact should decay back to zero. Averaging over all metaorders of size Q, one should therefore expect an apparent permanent impact given by:

$$\mathcal{I}_\infty(Q) = f(Q) \times I(Q) + (1 - f(Q)) \times I_R(Q), \tag{6.12}$$

where $f(Q)$ is the fraction of metaorders of volume Q that are truly informed and $I_R(Q)$ is the permanent, reactional part of impact – expected to be zero only if markets were truly efficient. A precise empirical determination of the size dependence of $I_R(Q)$ would be extremely interesting. However, this seems to be out of reach: not only would it require a large data set of metaorders reputed to be information-less (such as the portfolio transition trades of Ref. [11]), but also the error on the long-term asymptotic value of $I_R(Q)$ is bound to be very large, as Figure 6.3 shows. At this stage, it is thus difficult to confirm or infirm the validity of the theoretical arguments that predict a linear-in-Q dependence of $I_R(Q)$ [25, 69, 87].

Appendix

6.A More insights on the intraday price relaxation

After a buy (sell) metaorder, the order flow is on average biased towards buy orders (sell orders). On general grounds, it is natural to expect that the price dynamics of a metaorder reflects both its own impact and the one related to the order flow given by other metaorders executed by different market participants (or not). As discussed in Chapter 5, metaorders executed in the same time period and on the same asset tend to have similar signs, i.e. trading decisions are *crowded* since likely generated by similar trading signals. In fact, introducing the sign polarization parameter $m(\epsilon_i) \in [-1, 1]$ defined for each metaorder as follows

$$m(\epsilon_i) := \frac{\epsilon_i}{N-1} \sum_{k \neq i}^{N} \epsilon_k, \tag{6.13}$$

with N the daily number of metaorders per asset, we find that the metaorders are characterized by a positive average sign polarization $\mathbb{E}[m(\epsilon_i)] = 0.07$ (see the inset of the left panel in Figure 6.5).

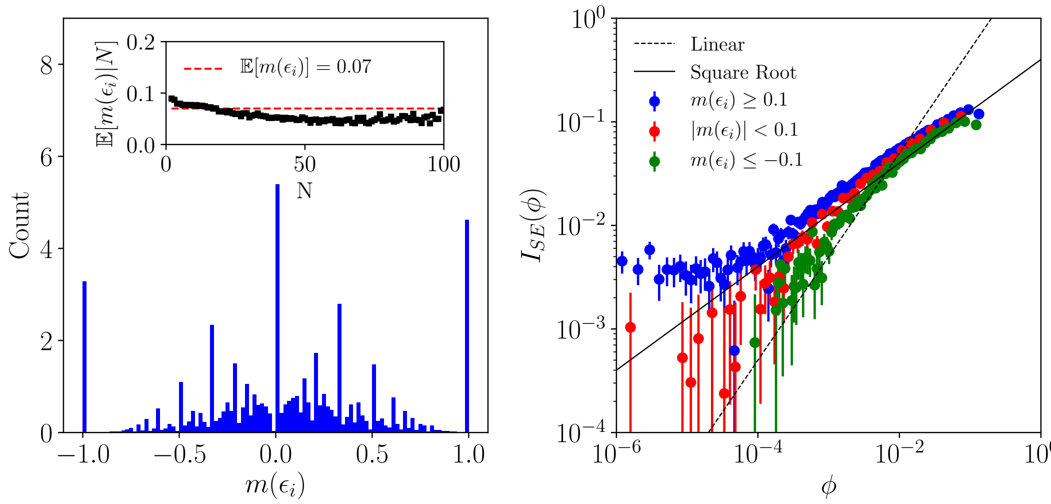

Figure 6.5: (Left panel) Empirical histogram of the sign polarization parameter $m(\epsilon_i)$ (see Eq. (6.13)): in the inset the conditional average of $m(\epsilon_i)$ respect to N is represented. (Right panel) Empirical Start-to-End market impact curves $I_{SE}(\phi)$ computed conditioning to several buckets of the polarization parameter $m(\epsilon_i)$: for a metaorder positively correlated with the intraday net order sign imbalance, i.e. $m(\epsilon_i) \geq 0.1$, the market impact is well described by a square-root law in the range $10^{-5} \lesssim \phi \lesssim 10^{-1}$.

It emerges that the Start-to-End price impact $I_{SE}(\phi)$ is statistically higher and well approximated by a square-root law when the executed metaorder is correlated in sign with the net order sign imbalance $\sum_{k \neq i} \epsilon_k$ given by the $N-1$ other *simultaneous* metaorders (see right panel of Figure 6.5). On the other hand, we find that this sign correlation implies a not-null plateau value for the intraday price decay as shown in the left panel of Figure 6.6 where conditioning to the $m(\epsilon_i)$ we obtain different price relaxation curves with distinct *asymptotic* values: if the metaorder is negatively correlated with the net order flow, i.e. $m(\epsilon_i) \leq -0.1$, then its decay goes to zero in less than one day while if it is positively correlated, i.e. $m(\epsilon_i) \geq 0.1$, we observe a permanent impact with a plateau around to the 2/3 of the peak impact. For these several regimes the propagator model, i.e. $\mathcal{I}_{\text{prop}}(z) = (1+z)^{1-\beta} - z^{1-\beta}$ with $z = V_{CE}/V_{SE}$, reproduces well the price relaxation curve with a β parameter opportunely fitted: to note that we find $\beta \simeq 0.5$ in the case of $m(\epsilon_i) \leq -0.1$. This is still valid conditioning the price decay to the participation rate η as shown in the right panel of Figure 6.6. In particular, we observe that the price relaxation is more pronounced for metaorders with large participation rate η: a possible explanation is that metaorders with lower participation rate revert more slowly as a consequence of the fact that the reversion process depends on the

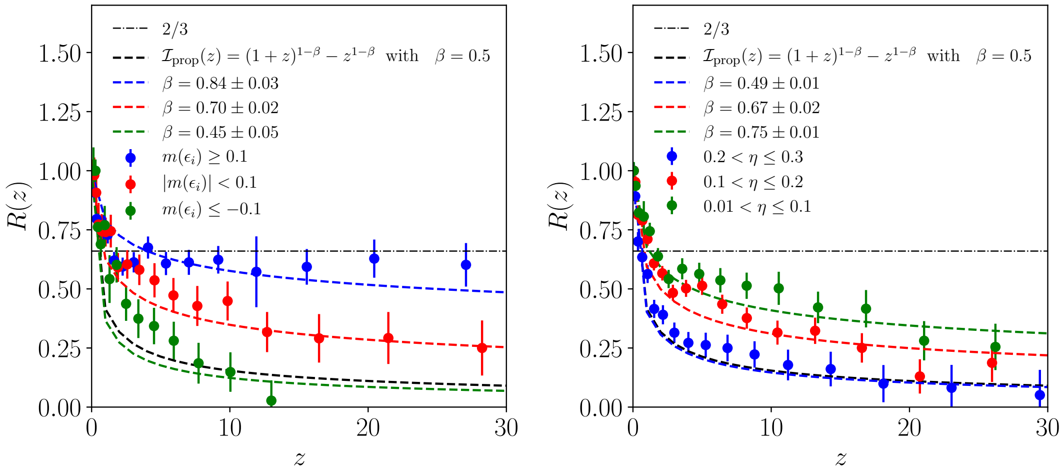

Figure 6.6: (Left panel) Intraday price relaxation curves in function of the ratio $z = V_{EC}/V_{SE}$ conditioned to the sign polarization parameter $m(\epsilon_i)$ defined in Eq. (6.13). (Right panel) Intraday price relaxation curves for different buckets in participation rate η. In both the panels the dashed lines represent respectively the best fit with the propagator model $\mathcal{I}_{\text{prop}}(z)$ for each sub-sample while the horizontal dashed line represents the 2/3-plateau predicted by the FGLW theory [15]. To note that an higher (lower) sign polarization $m(\epsilon_i)$ or participation rate η implies an higher (lower) asymptotic plateau for the price relaxation.

metaorder's detection by others market participants. In fact, the beginning or the end of a low participation rate metaorder is more difficult to detect, and it should require more time for a given level of certainty, giving a more sluggish reaction to completion of the metaorder.

Chapter 7

Are trading invariants really invariant? Trading costs matter

7.1 Introduction

Finding universal scaling laws between trading variables is highly valuable to make progress in our understanding of financial markets and market microstructure. Indeed, statistical physics has taught us that scaling laws between physical variables most often reflect the dynamics of complex scale-invariant systems, by that giving precious insights about the mechanisms underlying the phenomena at hand. Benoit Mandelbrot was the first to propose the idea of scaling in a financial and economic context [64], and ever since many have capitalised on his ideas, for a review see [65]. Relevant to this study, examples of universal laws between trading variables include the square root impact law of metaorders [2] or the relation between spread and volatility per trade [61]. Recently, Kyle and Obizhaeva posited an intriguing *trading invariance principle* that must be valid for a *bet*, theoretically defined as a sequence of orders with a fixed direction (buy or sell) belonging to a single trading idea [51, 52]. This principle supports the existence of a universal invariant quantity I_{ko} – expressed in dollars, independent of the asset and constant over time – which represents the average cost of a single *bet*. In particular, taking the share price S (in dollars per share), the square daily volatility σ_{d}^2 (in $\%^2$ per day), the total daily amount traded with *bets* V (shares per day) and the average unsigned volume of an individual *bet* Q (in shares) as relevant variables, dimensional analysis (see e.g. [66, 67]) suggests a relation of the form:

$$\frac{SQ}{I_{\mathrm{ko}}} = f\left(\sigma_{\mathrm{d}}^2 \frac{Q}{V}\right) \ , \tag{7.1}$$

where $f(\cdot)$ is a dimensionless function. Invoking the Modigliani-Miller capital structure irrelevance principle, which notably states that capital restructuring should always keep the product $S \times \sigma_{\mathrm{d}}$ while not affecting other variables, yields $f(x) \sim x^{-1/2}$ (see [51, 62] and Appendix 7.A). From Eq. (7.1) follows, up to a numerical factor, the 3/2-law:

$$I_{\mathrm{ko}} = \frac{\sigma_{\mathrm{d}} S Q^{3/2}}{V^{1/2}} := \frac{\mathcal{W}}{N^{3/2}} \,, \tag{7.2}$$

where $\mathcal{W} := \sigma_{\mathrm{d}} S V$ measures the total dollar amount of risk traded per day (also referred to as total exchanged risk or trading activity) while $N := V/Q$ represents the number of daily *bets* for a given financial instrument[1]. The above equation can be tested at different levels. First, I_{ko} is a random variable associated to each day and stock. The original (strong form of) trading invariance states that I_{ko} has invariant distribution to stock and time. A weaker, and more easily testable, form states that only the mean value of I_{ko} is invariant. Clearly, if this second form is violated (as shown empirically below), a *fortiori* the stronger form is violated. In other words, the 3/2-law can be interpreted with different degrees of universality[2] as discussed in [57]:

- *No universality*: the scaling relation $\mathcal{W} \sim N^{3/2}$ holds for some contracts and some time intervals τ over which $\mathcal{W}$ and N are computed. Than if the scaling law holds, the prefactor I_{ko} has a non-universal value dependent on the financial instrument and/or on τ.

- *Weak universality*: the 3/2-law holds for all contracts and some (possibly all) time intervals τ, but with a non-universal value of I_{ko}.

- *Strong universality*: the 3/2-law holds for all contracts and all time intervals τ, with a universal value of I_{ko}, independent of τ and of the contract type.

Let us stress that identifying an elementary *bet* in the market is not a straightforward task. Theoretically, a *bet* is defined as a trading idea typically executed in the market as many trades over several days. As suggested by Kyle and Obizhaeva in their original work [52], metaorders, i.e. a bundle of orders corresponding to a single trading decision typically traded incrementally through a sequence of child orders, can be considered a proxy of these *bets*[3]. Beyond the subtleties in the *bet*'s definition, there has been in the past few years empirical evidence that the scaling law discussed above matches patterns in financial data, at least approximately. The 3/2-law was empirically confirmed by Kyle and Obizhaeva using portfolio transition data related to rebalancing decisions made by institutional investors and executed by brokers [52]. Andersen et al. [56] reformulated suitably the *trading invariance hypothesis* at the single-trade level and showed that the equivalent version of Eq. (7.1) in such a setting holds remarkably well using public trade-by-trade data relative to the E-mini S&P 500 futures contracts. Benzaquen et al. [57] substantially extended these empirical results showing that the 3/2-law holds very precisely across 12 futures contracts and 300 single U.S. stocks, and across a wide

[1]See Appendix 7.A for more details on the derivation of the 3/2 scaling law.

[2]Note that here we only explore the daily level, time does not mean the same thing as in [57] where the authors varied the time intervals over which the variables were computed.

[3]In the following we will make use of such an approximation and use the words metaorder instead of *bet*.

range of time scales. Amongst others, Bowe et al. [55] examined market microstructure invariance relationships for equity markets using a subset of 25 equities from the FTSE 100 stocks traded on the London Stock Exchange, and Pohl et al. [58] provided additional empirical evidence that the intriguing 3/2-law holds on trades data from the NASDAQ stock exchange.

Notwithstanding, empirical data at the single transaction scale – see in particular [57] – revealed that while the 3/2-law is very robust, the invariant I_{ko} is actually quite far from invariant, as it varies from one asset to the other and across time, thus in favour of the *weak universality* degree. Note that this is consistent with the idea that a universal invariant with dollar units would be quite incongruous, given that the value of the dollar is itself stochastically time-dependent[4]. Benzaquen et al. [57] showed that a more suitable candidate for an invariant was actually the dimensionless $\mathcal{I} = I_{\mathrm{ko}}/\mathcal{C}$ where $\mathcal{C}$ denotes the spread trading costs.

Yet, single transactions are typically not the same as single *bets*. Large and medium sized orders are typically split in multiple transactions and traded incrementally over long periods of time. Public market data do not allow to infer the trading decision and to link different transactions to a single execution[5]. In order to test the trading invariance hypothesis at the metaorder level and its relation with trading costs, it is necessary to have a dataset of market-wide (i.e. not from a single institution) metaorders.

This is precisely the aim of the present chapter, which leverages on a heterogeneous dataset of metaorders extracted from the ANcerno database[6]. Although from a preliminary research Kyle and Fong found that proxies for *bet*s in ANcerno data have size patterns consistent with the proposed invariance hypothesis [51–54], to our knowledge such a thorough analysis at the metaorder level for a wide range of assets is still lacking.

Our main finding is that, while the scaling law $\mathcal{W} \sim N^{3/2}$ works surprisingly well independently of the chosen asset, the quantity I_{ko} is not invariant, as pointed out in [57] at the trade-by-trade level. In other words, for a given asset the 3/2-law (Eq. (7.2)) holds, but the invariance principle implying that I_{ko} is the same for all assets does not. We show that the latter quantity is strongly correlated with transaction costs, including spread and impact costs. This leads us to introduce new invariants, obtained by dividing I_{ko} by the trading costs, and which appear to fluctuate very little across stocks. Finally we show that the observed small dispersion of the new invariants can be connected with three microstructural properties: (i) the linear relation between spread and volatility per transaction; (ii) the near invariance of the metaorder size distribution, and (iii) of the total volume and number fractions of the metaorders across different stocks.

[4] Note that Kyle and Obizhaeva commented on how to modify their invariance principle in an international context. In particular they suggested that "invariance relationships can also be applied to an international context in which markets have different currencies or different real exchange rates" by scaling to "the nominal cost of financial services calculated from the productivity-adjusted wages of finance professionals in the local currency of the given market during the given time period" [51].

[5] In fact, for example, Kyle and Obizhaeva tackled this problem investigating a proprietary dataset of portfolio transitions [52].

[6] In Ref. [52] the authors claim that the ANcerno database includes more orders than the dataset of portfolio transitions they used in their work.

The chapter is organised as follows. In Section 7.2 we describe the dataset collecting trading decisions of institutional investors operating in the U.S. equity market. In Section 7.3 we show that the 3/2-law holds surprisingly well at the daily level independently of the time period, of the market capitalisation and of the economic sector. In Section 7.4 we compute the invariant I_{ko} and we argue in favour of *weak universality*. We propose a more natural definition for a trading invariant that accounts both for the spread and the market impact costs and we exhibit the microstructural origin of its small dispersion. Some conclusions and open questions are presented in Section 7.5.

7.2 Data

Our analysis relies on a database made available by ANcerno, a leading transaction-cost analysis provider. Our dataset counts heterogeneous institutional investors placing large buy or sell orders executed by a broker as a succession of smaller orders belonging to the same trading decision of a single investor (for more details see Chapter 3). Our sample includes the period January 2007 – June 2010 for a total of 880 trading days. Only metaorders completed within at most a single trading day are held. Further, we select stocks belonging to the Russell 3000 index, thereby retaining $\sim$ 8 million metaorders distributed quite uniformly in time and representing $\sim 5\%$ of the total reported market volume, regardless of market capitalisation (large, mid and small) and economical sectors (basic materials, communications, consumer cyclical and non-cyclical, energy, financial, industrial, technology and utilities). As can be seen in [10,35,41], which use very similar filtering of the dataset, the distribution of metaorder duration, traded volume, and participation rate are very heterogeneous, spanning several orders of magnitude. When considering the number N of metaorders traded in a day for a stock, the left panel of Figure 7.8 in Appendix 7.C shows a quite heterogeneous distribution with on average approximately 5 metaorders executed per day in each asset.

7.3 The 3/2-law

Here we investigate the *trading invariance hypothesis* at the daily level. The daily timescale choice avoids an elaborate analysis of when precisely each metaorder starts and ends, thereby averaging out all the non-trivial problems related to the daily simultaneous metaorders executed on the same asset.

7.3.1 Exchanged risk

From the metaorders executed on the same stock during the same day we compute the total exchanged volume in dollars: $\sum_{i=1}^{N} S_i Q_i$, where N is the number of daily metaorders per asset in the ANcerno database, Q_i and S_i are respectively the number of shares and the volume weighted average price of the i-th available metaorder. We then define the total daily exchanged ANcerno risk per asset as:

$$\mathcal{W} := \sum_{i=1}^{N} \mathcal{W}_i \ , \quad \text{with} \quad \mathcal{W}_i := \sigma_{\rm d} Q_i S_i \ , \tag{7.3}$$

and where $\sigma_{\rm d}$ denotes the daily volatility per asset, computed as $\sigma_{\rm d} = (S_{\rm high} - S_{\rm low})/S_{\rm open}$ from the high, low, and open daily prices only[7]. The statistical properties of the metaorders, in terms of their associated risk $\mathcal{W}_i$ and of their total daily number N per asset are discussed in Appendix 7.C. It is found that the empirical distributions of the traded risks $\mathcal{W}_i$ and $\mathcal{W}$ span almost eight orders of magnitude. This is important because a careful testing of the scaling relation predicted by the trading invariance hypothesis requires a large variability of the considered variables. Thus the ANcerno database is ideal for this testing exercise.

7.3.2 Empirical evidence

We introduce the mean daily exchanged risk $\langle \mathcal{W} \rangle_N$ conditional to the number of metaorders N. This is empirically estimated by the quantity:

$$\langle \mathcal{W} \rangle_N := \frac{\sum_{\ell: N^{(\ell)}=N} \mathcal{W}^{(\ell)}}{\sum_{\ell: N^{(\ell)}=N} 1} \ , \tag{7.4}$$

where for each day ℓ the total daily exchanged risk is given by $\mathcal{W}^{(\ell)} := \sum_{i=1}^{N^{(\ell)}} \mathcal{W}_i^{(\ell)}$ with $N^{(\ell)}$ the number of daily metaorder per asset and $\mathcal{W}_i^{(\ell)} := \sigma_{\rm d}^{(\ell)} Q_i^{(\ell)} S_i^{(\ell)}$. To test the 3/2-law we bin the data (one observation per stock per day) depending on N and we plot it against $\langle \mathcal{W} \rangle_N$ in log-log scale. We consider different subsets of stocks, depending on market capitalisation, economical sector, investigated period, and we perform the linear regression of $\log\langle \mathcal{W} \rangle_N$ versus $\log N$.

As shown in the first three panels of Figure 7.1 the scaling $\langle \mathcal{W} \rangle_N \sim N^{3/2}$ holds well independently of the conditioning to market capitalisation, economical sector, and time period. The insets show the estimated exponent which is always quite close to 3/2. Slight deviations may have different origins but can mostly be attributed to the heterogeneous sample's composition in terms of stocks for each bucket in N. The 3/2-law is also valid for individual stocks, as shown in the bottom right panel of Figure 7.1, where data from 10 randomly chosen stocks are displayed. We perform the above regression on a larger sub-sample of 200 stocks randomly chosen and the histogram of the slopes (exponents), shown in the bottom left inset, is well centered around the 3/2-value. This shows that the 3/2-exponent works very well in describing the scaling relation between $\langle \mathcal{W} \rangle_N \sim N^{3/2}$ and N.

[7] We checked that the results discussed in this chapter are still valid using other definitions of the daily volatility and of the price in analogy to what done for example in [52]. Specifically, the results are still valid when computing $\sigma_{\rm d}$ with the Rogers-Satchell volatility estimator [60] or as the monthly averaged daily volatility, i.e. $\bar{\sigma}_{\rm d} = \sum_{\ell=1}^{25} \sigma_{d,\ell}$ and/or defining the price S_i as the closing price of the day before the metaorder's execution.

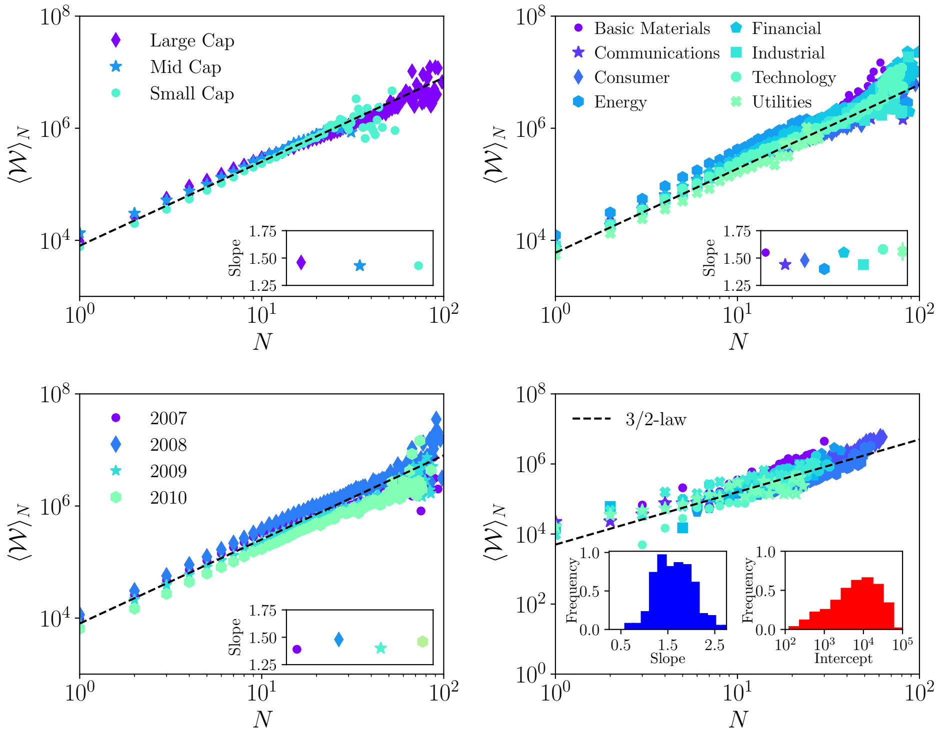

Figure 7.1: Mean daily exchanged risk $\langle \mathcal{W} \rangle_N$ conditional on the daily number N of metaorders per asset for different market capitalisation (top left panel), the economical sector (top right panel), and the time period (bottom left panel). The insets show the slopes obtained from linear regression of the data, firstly averaged with respect to N and secondly log-transformed. The bottom right panel shows a plot of $\langle \mathcal{W} \rangle_N$ as function of N for a subset of 10 stocks chosen randomly from the pool of around three thousand U.S. stocks: the two insets represent respectively the empirical distribution of the slopes and of the y-intercept, i.e. $\langle I_{\text{ko}} \rangle = \langle \mathcal{W} \rangle_N / N^{3/2}$, obtained from linear regression of a larger sub-sample of 200 stocks randomly chosen, firstly averaged with respect to N and secondly log-transformed considering each stock separately.

The bottom right inset in the bottom right panel of Figure 7.1 shows the histogram of the intercept $\langle I_{\text{ko}} \rangle$ obtained from the regression $\log \mathcal{W} = \log \langle I_{\text{ko}} \rangle + \beta \log N$ done for individual stocks and using the binned data as shown in the main panel. It is evident that there is a very large dispersion (note that the abscissa is in log scale), which indicates that I_{ko} in Eq. (7.2) is not constant across different stocks. More empirical insights on the origin of the 3/2-law are presented in Appendix 7.B.

7.4 The trading invariant

The conjecture that the quantity $\langle I_{\rm ko}\rangle = \langle \mathcal{W}\rangle_N / N^{3/2}$ is invariant across different contracts is clearly rejected by the empirical analysis performed in the previous section. Indeed, the quantity $\langle I_{\rm ko}\rangle$ varies by at least one order of magnitude across different stocks. This result goes against the *strong universality* version of the *trading invariance hypothesis* which states that both the average value $\langle I_{\rm ko}\rangle$ and the full probability distribution of $I_{\rm ko} = \mathcal{W}/N^{3/2}$ should be invariant across products. Dimensionally $I_{\rm ko}$ is a cost (i.e. it is measured in dollars) and indeed the *trading invariance hypothesis* posits that the cost of a metaorder is invariant. Using the identification of metaorders and *bets* we can use the ANcerno dataset to estimate the trading cost, including a spread and a market impact component. Below, we show that $I_{\rm ko}$ and the trading costs are highly correlated, and therefore we propose new invariants based on their ratio.

7.4.1 Trading costs and trading invariants

Trading costs are typically divided into fees/commissions, spread, and market impact. For large orders, like those investigated here, fees/commissions typically account for a very small fraction and therefore we will neglect them. We shall however take into consideration both the spread cost (as was done at the single-trade level in [57]) and the market impact cost as computed from the well established square-root law (see e.g. [4, 10, 12, 35, 41, 46, 50]). We thus define the average daily metaorder's trading cost as:

$$\mathcal{C} = \mathcal{C}_{\rm spd} + \mathcal{C}_{\rm imp} = Y_{\rm spd} \times \frac{1}{N}\sum_{i=1}^{N} \mathcal{S} Q_i + Y_{\rm imp} \times \frac{1}{N}\sum_{i=1}^{N} \sigma_{\rm d} Q_i S_i \sqrt{\frac{Q_i}{V_{\rm d}}} = Y_{\rm spd} \times \mathcal{C}^0_{\rm spd} + Y_{\rm imp} \times \mathcal{C}^0_{\rm imp}\,, \tag{7.5}$$

with $\mathcal{S}$ the average daily spread[8], $V_{\rm d}$ the total daily market volume, and $Y_{\rm spd}, Y_{\rm imp}$ two constants to be determined. The factor $Y_{\rm spd}$ depends, among other things, on the fraction of trades of the metaorder executed with market orders, whereas $Y_{\rm imp}$ only weakly depends on the execution algorithm and is typically estimated to be very close to unity [2, 10, 34]. Thus, while $\mathcal{C}_{\rm imp}$ is a quite faithful estimation of the impact cost of the metaorders in a day and stock, $\mathcal{C}_{\rm spd}$ is an upper bound, reached if all the considered metaorders are executed with market orders.

The empirical properties of $\mathcal{C}_{\rm spd}$ and $\mathcal{C}_{\rm imp}$ and the relative importance of the two terms as a function of the metaorder size are illustrated in Appendix 7.C. As expected, at the single metaorder level, spread cost is dominant for small metaorders, while market impact cost is dominant for large ones. At the aggregated level, the average daily market impact cost $\mathcal{C}_{\rm imp}$ accounts on average for approximately half of the total daily trading average cost.

To determine $Y_{\rm spd}$ and $Y_{\rm imp}$ we perform an ordinary least square regression of the KO invariant $I_{\rm ko}$ with respect to the daily average cost $\mathcal{C}$ defined for each asset by Eq. (7.5).

[8]The daily spread is not provided in the ANcerno dataset: we computed it as the time average spread across the day using public available market data.

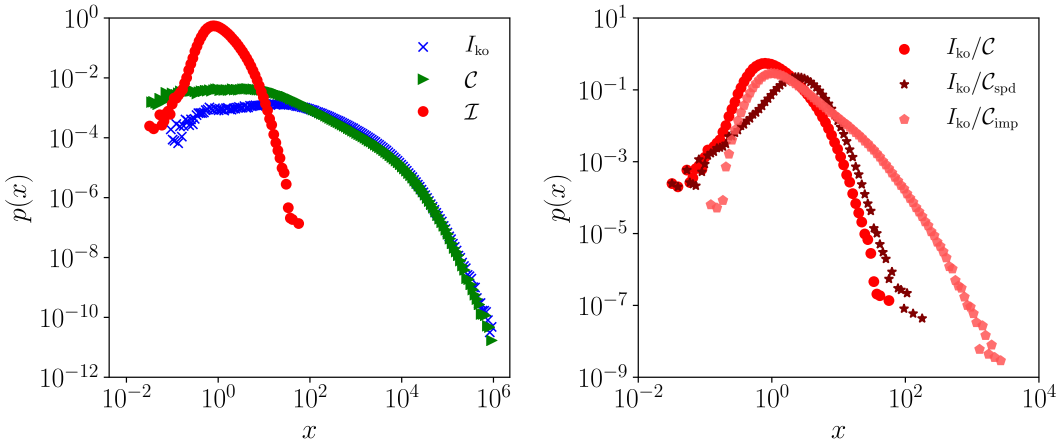

Figure 7.2: (Left panel) Empirical distributions of the KO invariant $I_{\rm ko} = \mathcal{W}/N^{3/2}$, of the daily average metaorder's total trading cost $\mathcal{C}$ (using $Y_{\rm spd} = 3.5$ and $Y_{\rm imp} = 1.5$ in Eq. (7.5)), and of the dimensionless invariant $\mathcal{I} = I_{\rm ko}/\mathcal{C}$. (Right panel) Empirical distributions in log-log scale of the KO invariant $I_{\rm ko}$ rescaled respectively by the total daily average cost $\mathcal{C}$, by the spread cost $\mathcal{C}_{\rm spd}$ and by the market impact cost $\mathcal{C}_{\rm imp}$.

We obtain $Y_{\rm spd} \simeq 3.5 \pm 0.2$, $Y_{\rm imp} \simeq 1.5 \pm 0.1$ and a coefficient of determination $r^2 \simeq 0.8$. These results show that the original KO invariant is indeed strongly correlated with the trading cost. Since these costs have no a priori reason to be universal, this explains why $I_{\rm ko}$ is not invariant.

Guided by such results and by the fact that a market microstructure invariant, if any, should be dimensionless, we define new invariants by dividing the original KO invariant $I_{\rm ko}$ by the cost of trading. Therefore, we consider three different specifications, namely:

$$\mathcal{I} := \frac{I_{\rm ko}}{\mathcal{C}} \,, \qquad \mathcal{I}_{\rm spd} := \frac{I_{\rm ko}}{\mathcal{C}_{\rm spd}} \,, \qquad \mathcal{I}_{\rm imp} := \frac{I_{\rm ko}}{\mathcal{C}_{\rm imp}} \,. \tag{7.6}$$

The left panel of Figure 7.2 shows the empirical distribution of the original KO invariant $I_{\rm ko}$ together with that of $\mathcal{I}$, and of the total trading cost $\mathcal{C}$. It is visually quite clear that rescaling by the cost dramatically reduces the dispersion, and that the distribution of $I_{\rm ko}$ is very similar to that of $\mathcal{C}$, despite some deviation for small value. The right panel compares the distribution of $\mathcal{I}$ with that of the other two new invariants. A quantitative comparison is provided in Table 7.1, which reports the mean, the standard deviation, the coefficient of variation[9] (CV) of $I_{\rm ko}$ and of the three new invariants. It is clear that, due to the correlation between $I_{\rm ko}$ and $\mathcal{C}$, the new invariants $\mathcal{I}_{\rm spd}$ and $\mathcal{I}_{\rm imp}$ have a much smaller CV than $I_{\rm ko}$. Since the distributions have clear fat tails, we also

[9] The coefficient of variation (CV) is the ratio of standard deviation and mean, an indicator of distribution peakedness.

implemented the Gini coefficient, as done in [58]. The table indicates that also in this case the new invariants are much more peaked than I_{ko}.

Table 7.1: Statistics of the different invariants, namely the original KO invariant I_{ko} (left), and the three new ones rescaled by cost (right). CV stands for coefficient of variation.

	$I_{\mathrm{ko}} \cdot 10^3$ (\$)	$\mathcal{I}$	$\mathcal{I}_{\mathrm{spd}}$	$\mathcal{I}_{\mathrm{imp}}$
mean	6.33	2.20	4.70	7.8
st. dev.	11	1.84	3.11	12.2
CV	1.74	0.84	0.66	1.56
Gini Coefficient	0.77	0.33	0.39	0.65

7.4.2 Origin of the small dispersion of the new invariants

Here we investigate the origin of the small dispersion of the new invariants. Let us first consider only the market impact cost normalisation and rewrite $\mathcal{I}_{\mathrm{imp}}$ with the approximation of $S_i \simeq S$ for all the metaorders executed in a day and on the same stock as:

$$\mathcal{I}_{\mathrm{imp}} = \frac{N \sum_{i=1}^{N} \sigma_{\mathrm{d}} S_i Q_i}{Y_{\mathrm{imp}} N^{3/2} (\sigma_{\mathrm{d}} \sum_{i=1}^{N} S_i Q_i \sqrt{Q_i/V_{\mathrm{d}}})} = \frac{1}{Y_{\mathrm{imp}} \sqrt{\upsilon}} \frac{[Q]^{3/2}}{[Q^{3/2}]} = \frac{1}{Y_{\mathrm{imp}} m \sqrt{\upsilon}} \,, \tag{7.7}$$

where $\upsilon := V/V_{\mathrm{d}}$ with $V := \sum_{i=1}^{N} Q_i$ the total ANcerno metaorder volume, $[\cdot]$ a daily average operation per stock, i.e. $[x] := \frac{1}{N} \sum_{i=1}^{N} x_i$, and $m > 1$ the normalised 3/2-th moment of the number of shares of a metaorder, which depends on the shape of the distribution of metaorder size. We have checked that m as well as υ are, to a first approximation, independent of the stock (see left and central panels in Figure 7.3) indicating that the distribution of metaorder size is, to a large degree, universal and that the ANcerno database is representative of the trading across all stocks. These observations explains why $\mathcal{I}_{\mathrm{imp}}$ is also, to a large degree, stock independent.

For the total cost normalisation, our understanding of the invariance property relies on the following empirical fact. The average spread is proportional to the volatility per trade, that is $\mathcal{S} = c\, S \sigma_{\mathrm{d}} / \sqrt{N_{\mathrm{d}}}$, where N_{d} is the total number of daily transactions per asset and c is a stock independent numerical constant, see [34, 61]. Indeed, the above arguments taken together show that the dimensionless quantity $\mathcal{I}$ can be written as:

$$\mathcal{I} = \frac{1}{Y_{\mathrm{spd}} c \sqrt{n} + Y_{\mathrm{imp}} m \sqrt{\upsilon}}, \tag{7.8}$$

where $n := N/N_{\mathrm{d}}$ is found to be stock independent (see right panel in Figure 7.3). Therefore $\mathcal{I}$ is also stock independent. Finally, the fact that the CV of $\mathcal{I}$ is less than both that of $\mathcal{I}_{\mathrm{spd}}$ and $\mathcal{I}_{\mathrm{imp}}$ suggests that KO's invariant is commensurate to the total cost of trading, including both the spread cost and the impact cost.

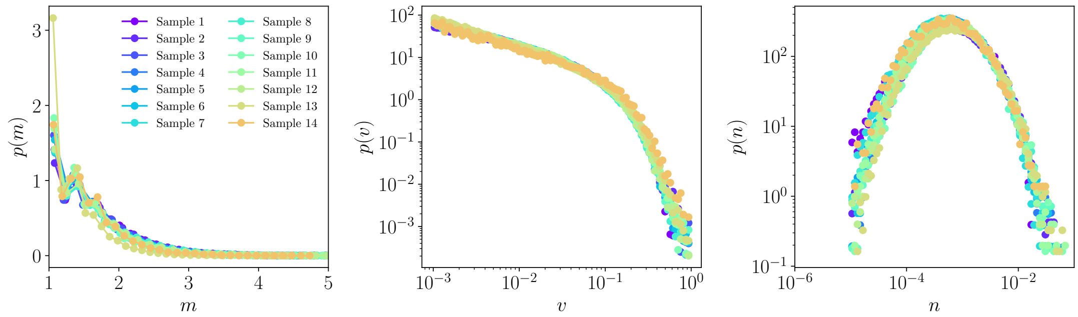

Figure 7.3: Empirical distributions of the ratios $m = [Q^{3/2}]/[Q]^{3/2}$ (left panel), $v = V/V_\mathrm{d}$ (central panel) and $n = N/N_\mathrm{d}$ (right panel), all three computed at the daily level for each asset: we randomly group the stocks in equally sized samples and for each of them we compute the empirical distribution respectively of m, v and n finding that they are, to a first approximation, stock independent.

7.5 Conclusions

In this chapter we empirically investigated the *market microstructure invariance hypothesis* recently proposed by Kyle and Obizhaeva [51, 52]. Their conjecture is that the expected dollar cost of executing a *bet* is constant across assets and time. The ANcerno dataset provides a unique laboratory to test this intriguing hypothesis through its available metaorders which can be treated as a proxy for *bet*s, i.e. a decision to buy or sell a quantity of institutional size generated by a specific trading idea. Let us summarise what achieved in this chapter:

- Using metaorders issued for around three thousand stocks, we showed that, at the daily timescale interval, the 3/2-law between exchanged risk $\mathcal{W}$ and number of metaorders N is observed independently of the year, the economical sector and the market capitalisation.

- The trading invariant $I_\mathrm{ko} = \mathcal{W}/N^{3/2}$ proposed by Kyle and Obizhaeva is non-universal: both its average value $\langle I_\mathrm{ko} \rangle$ and its distribution clearly depend on the considered stocks, in favour of a *weak universality* interpretation. Furthermore, this quantity has dollar units which makes its hypothesised invariance rather implausible.

- On the basis of dimensional and empirical arguments, we propose a dimensionless invariant defined as a ratio of I_ko and of the metaorder's total cost, which includes both spread and market impact costs. We find a variance reduction of more than 50%, qualitatively traceable to the proportionality between spread and volatility per trade, and the near invariance of the distributions of metaorder size, of the volume fraction and number fraction of metaorders across stocks.

Our empirical analysis has allowed to show that the *trading invariance hypothesis* holds at the metaorder level in a strong sense provided one considers the exchanged risk and the total trading cost of the metaorders. This is in the spirit of Kyle and Obizhaeva's arguments, but takes into account the fact that transaction costs are both asset and epoch dependent. As anticipated in [57], our results strongly suggest that *trading invariance* is a consequence of the validity of the square-root law for market impact as well as to the proportionality between spread and volatility as discussed in [57,61,63]. More generally, we interpret it as a result of the endogeneisation of costs in the trading decision of market participants. This is an alternative explanation with respect to the one invoking the Modigliani-Miller theorem and proposed by Kyle and Obizhaeva. It would actually be quite interesting to investigate other markets such as bond, currency or futures markets, for which the Modigliani-Miller theorem is totally irrelevant, while *trading invariance* still holds – at least at the level of single trades [56, 57]. Finally, note that differences in market structure across countries, such as execution mechanisms, fees and regulations could also challenge the validity of the results presented here.

Appendix

7.A Trading invariance principle: dimensional analysis + leverage neutrality

Let us describe the guide lines for the derivation of the *trading invariance principle* stated in the work of Kyle and Obizhaeva [51]. From a general point of view Kyle and Obizhaeva (KO) develop their conjecture applying *dimensional analysis* combined with the *leverage neutrality principle* related to the Modigliani-Miller theorem [62].

Their idea to invoke dimensional analysis follows from the intuitive argument that a meaningful relation between quantities involving some dimensions should not be affected by the units in which these dimensions are measured. In their theoretical framework of the trading invariance principle it is assumed that the relevant dimensions are represented by time $\mathbb{T}$, shares $\mathbb{S}$, and money $\mathbb{U}$. Based on these dimensions, Kyle and Obizhaeva turn to the idea of dimensional analysis from which the validity of a considered relation should not depend on whether we measure time in seconds or minutes, shares in single shares or in packages of hundred shares, and money in Dollars or in Dollars-cents. From a formal point of view, this means that a function $g(\cdot)$ relating a quantity of interest Y to a set of explanatory variables $X_1, \cdots, X_n$

$$Y = g(X_1, \cdots, X_n) \tag{7.9}$$

is dimensional invariant if it is invariant under the rescaling of the involved dimensions (in this case represented by $\mathbb{T}$, $\mathbb{S}$, and $\mathbb{U}$). In the trading invariance framework the quantity of interest is the arrival rate of bet for a given stock given by

- $N = N_t^{t+\tau}$, the number of bet within a time interval $[t, t+\tau]$ and with dimensional unit equals to $[N] = \mathbb{T}^{-1}$.

On the other side the explanatory variables which can determine the number of bets N in a given time interval $[t, t+\tau]$ are

- the traded volume of the stock $V = V_t^{t+\tau}$ during the time interval $[t, t+\tau]$, measured in units of shares per time $[V] = \mathbb{S}/\mathbb{T}$;
- the average price of the stock $S = S_t^{t+\tau}$ in the time interval $[t, t+\tau]$, measured in units of money per share $[S] = \mathbb{U}/\mathbb{S}$;
- the square variance of the stock $\sigma^2 = (\sigma^2)_t^{t+\tau}$ in the time interval $[t, t+\tau]$, with dimensional unit equals to $[\sigma^2] = \mathbb{T}^{-1}$.

From these assumptions Kyle and Obizhaeva postulate that the number of bets N per time interval τ can be fully explained from the following four dimensional quantities - the price S per share (in dollars), the square volatility σ^2 (in $\%^2$ per time interval τ), the total amount traded with bets V (shares per time interval τ), and the average cost of a single bet I_{ko} representing the trading invariant cost (in dollars) - through a function $h(\cdot)$, i.e.

$$N = \sigma^2 \, h\left(\frac{SV}{\sigma^2 I_{\text{ko}}}\right), \tag{7.10}$$

where the four explanatory variables above are combined taking into account the invariance relations pertaining their physical dimensions $\mathbb{S}$, $\mathbb{T}$, and $\mathbb{U}$. However, to determine the shape of the unknown function $h(\cdot)$ Kyle and Obizhaeva introduce the *leverage neutrality principle* [68] which captures the intuition of the Modigliani-Miller theorem stating that a firm mix of equity and risk free debt securities does not affect the value of a firm. Defining the leverage quantity as

$$L = \frac{\text{total assets}}{\text{equity}} \tag{7.11}$$

it follows that multiplying L by a factor $A > 1$ is equivalent to paying out $(1 - A^{-1})$ of the equity as cash-dividends; on the other side, multiplying L by a factor $0 < A < 1$ corresponds to raising new capital in order to increase the firm equity by a factor A^{-1}. Then from this *leverage neutrality principle* it emerges that varying the leverage L by a factor A

- S changes by a factor A^{-1} and
- σ^2 changes by a factor A^2.

For example, setting $A = 2$ corresponds to paying out half of the equity as dividends so that each share yields a dividend of $(1 - A^{-1})S = S/2$. The stock price is thus multiplied by $A^{-1} = 1/2$ while the volatility σ is multiplied by $A = 2$, and the remaining quantities are not affected by changing the leverage, in accordance with the insight of the Modigliani-Miller theorem. The economic reason is that the value of the assets of the corresponding company and hence the associated risk does not change: in other words the capital restructuring between debt and equity should keep the product $\sigma \times S$ constant without affecting the other variables.

Applying the leverage neutrality principle to Eq. (7.10), which implies its invariance with respect to the scale transformation $S \to A^{-1}S$ and $\sigma^2 \to \sigma^2 A^2$, Kyle and Obizhaeva demonstrate through the Vaschy-Buckingham π-theorem (see [117] for reference) that $h(x) \sim x^{2/3}$ in Eq. (7.10), or equivalently $f(x) \sim x^{-1/2}$ in Eq. (7.1). From this result it follows finally the 3/2-law which states that the KO trading invariant cost is given by

$$I_{\text{ko}} = \frac{\sigma S Q^{3/2}}{V^{1/2}} = \frac{\mathcal{W}}{N^{3/2}}, \tag{7.12}$$

with $\mathcal{W} = \sigma S V$ the total risk exchanged with bets and $Q = V/N$ the average size of a bet executed on a time interval τ. Note that in this chapter we investigated the trading invariance principle at the daily time scale fixing τ=1 day and $\sigma = \sigma_{\text{d}}$.

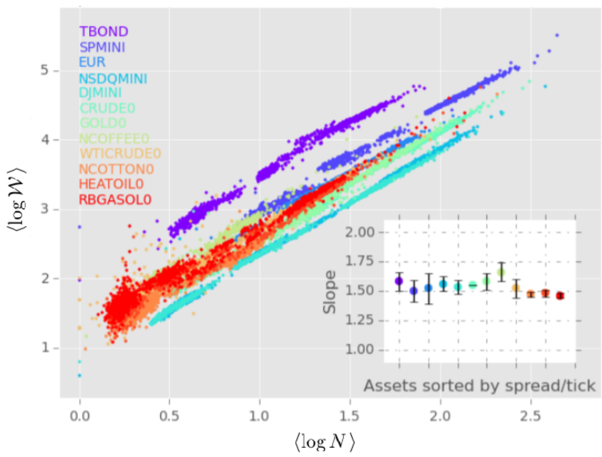

Figure 7.4: Scatter plot of $\langle \log \mathcal{W} \rangle$ vs. $\langle \log N \rangle$ for twelve different future contracts spanning over three years, from January 2012 to December 2014, and sorted by spread over tick values from cold (large ticks) to warm colours (small ticks) (figure reproduced from [57]). To test the trading invariance hypothesis the authors firstly average over all days the logarithm of the considered quantities for each fixed $\tau = 1$ minute bin and secondly take the average over each these bins. The inset shows the slopes obtained from linear regression of $\log\mathcal{W}$ vs. $\log N$, which are all clustered around 3/2 confirming therefore the 3/2-law at the intraday time scale.

At this point let us review the recent empirical evidences discussed in literature of the trading invariance principle and then of the 3/2-law. Firstly, the 3/2-law was empirically confirmed by Kyle and Obizhaeva using portfolio transition data [52]. Portfolio transitions correspond to rebalancing decisions by institutional investors and executed by brokers. However, these trades only reflect part of the market activity, and it is furthermore not obvious that these portfolio transitions can be associated with bets. To overcome this issue Andersen et al. [56] reformulated the KO trading invariance principle in a way that can be tested on public trade level data. Their analysis on the E-mini S&P 500 futures contracts showed that the 3/2-law holds remarkably well at the single-trade level. In this context, Q denotes the average volume of trades and N is the total number of trades executed within a time interval τ (fixed to 1 minute in their analysis).

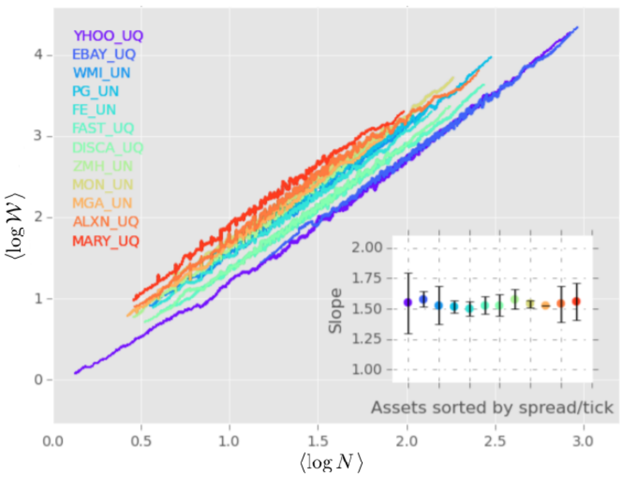

Figure 7.5: Centred rolling average with window size=100 of the scatter plot of $\langle \log \mathcal{W} \rangle$ vs. $\langle \log N \rangle$ for a random subset of twelve different stocks chosen from a pool of three hundred U.S. stocks and sorted by spread over tick values from cold (large ticks) to warm colour (small ticks) (figure reproduced from [57]). The empirical analysis is done considering $\tau = 5$ minutes bins using trades and quotes data from January 2012 to December 2012, extracted from the primary market of each stock (NYSE/NASDAQ). The inset shows the slopes obtained from linear regression of $\log \mathcal{W}$ vs. $\log N$, which are clearly clustered around the 3/2 value.

Benzaquen et al. [57] address the same question by investigating twelve additional futures contracts as well as three hundred U.S. stocks. Aiming to confirm that $\beta = 3/2$ in the scaling $\mathcal{W} \sim N^{\beta}$, where $\mathcal{W}$ is the total exchanged risk through trades executed in the time interval τ, they estimate β singularly for each stock (see Figure 7.4) and future (see Figure 7.5) across a wide range of time scales τ, finding that in average $\beta = 1.54 \pm 0.11$. Moreover, they show that the distribution of the trading invariant I_{ko} depends significantly on the studied asset and thus conclude that the 3/2-law holds only with weak universality. As an additional contribution, the authors reveal that the inclusion of the average spread cost per trade is beneficial in the sense that their proposed invariant is almost constant for different assets.

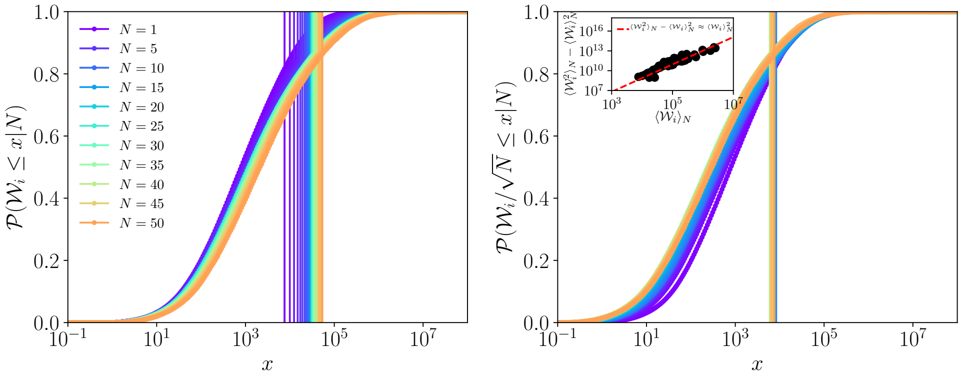

Figure 7.6: Empirical cumulative distribution of the traded metaorder's risk $\mathcal{W}_i = \sigma_{\mathrm{d}} Q_i\, S_i$ without (left panel) and with (right panel) rescaling by the square root of the daily number N of metaorders per asset. The colored vertical lines represent the location of the average for each sample conditional on N. To note that also if the empirical distribution is not an invariant function of N, we observe that $\langle \mathcal{W}_i/\sqrt{N}\rangle_N \simeq$ const., as evident from the vertical lines in the right panel, which is at the origin of the measured 3/2-law. Furthermore, as shown in the inset the variance $\langle \mathcal{W}_i^2\rangle_N - \langle \mathcal{W}_i^2\rangle_N$ scales linearly with N, i.e. $\langle \mathcal{W}_i^2\rangle_N - \langle \mathcal{W}_i^2\rangle_N \approx \langle \mathcal{W}_i\rangle_N^2 \sim N$.

7.B The 3/2-law under the microscope

One may rightfully wonder whether it is possible to understand the 3/2-law from the statistical properties of the metaorders. For this purpose we start by investigating the individual metaorder's risk $\mathcal{W}_i$ distribution properties as a function of N. We find that when rescaling the metaorder's risk $\mathcal{W}_i$ by the square root of the number N of daily metaorders per asset one obtains a conditional cumulative distribution $\mathcal{P}(\mathcal{W}_i/\sqrt{N}|N)$ dependent on N but with a mean $\langle \mathcal{W}_i/\sqrt{N}\rangle_N$ invariant on N (see Figure 7.6)[10]. It emerges then that the conditional average metaorder risk $\mathcal{W}_i$ can be predicted from the number N of daily metaorders per asset since $\langle \mathcal{W}_i\rangle_N$ scales as N^γ with $\gamma \simeq 0.5$, that is $\langle \mathcal{W}_i\rangle_N \sim \sqrt{N}$ [11]. It immediately follows that combining this empirical result and the linearity property of the mean, one recovers the 3/2-law $\langle \mathcal{W}\rangle_N \sim N^{3/2}$, since:

$$\langle \mathcal{W}\rangle_N = \langle \sum_{i=1}^{N} \mathcal{W}_i\rangle_N = \sum_{i=1}^{N} \langle \mathcal{W}_i\rangle_N = N\, \langle \mathcal{W}_i\rangle_N \sim N\sqrt{N} = N^{3/2}. \tag{7.13}$$

To explain the scaling $\langle \mathcal{W}_i\rangle_N \sim \sqrt{N}$ through the product $\langle \sigma_{\mathrm{d}}\rangle_N \times \langle Q_i S_i\rangle_N$ we need to check for the correlation between the daily volatility σ_{d} and the volume in dollars $Q_i S_i$ of a metaorder, which is found to be $\langle \mathbb{C}(\sigma_{\mathrm{d}}, Q_i S_i)\rangle \approx 3\times 10^{-2}$, where the average $\langle\cdot\rangle$ is done

[10] Here $\langle\cdot\rangle$ denotes the average over all days and stocks present in the sample.

[11] In analogy, the variance $\langle \mathcal{W}_i^2\rangle_N - \langle \mathcal{W}_i^2\rangle_N$ scales linearly with N, i.e. $\langle \mathcal{W}_i^2\rangle_N - \langle \mathcal{W}_i^2\rangle_N \approx \langle \mathcal{W}_i\rangle_N^2 \sim N$.

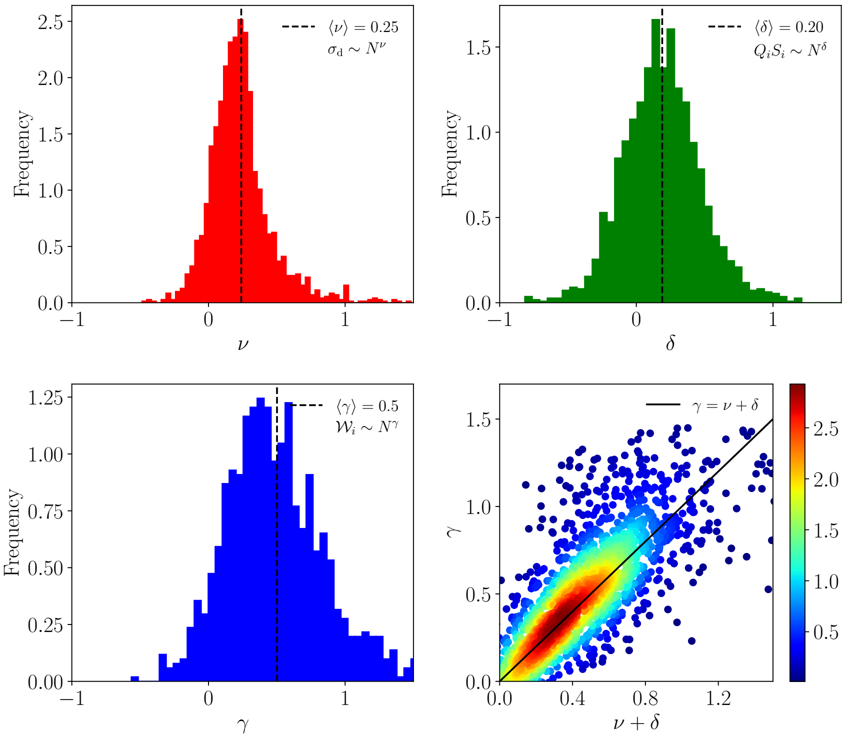

Figure 7.7: (Top left panel) Empirical distribution of the scaling exponent ν computed for each stock regressing $\sigma_{\mathrm{d}} \sim N^{\nu}$: in average $\langle \nu \rangle = 0.25$ as shown by the dashed black line. (Top right panel) Empirical distribution of the scaling exponent δ computed for each stock regressing $Q_i S_i \sim N^{\delta}$: in average $\langle \delta \rangle = 0.20$ as shown by the dashed black line. (Bottom left panel) Empirical distribution of the scaling exponent γ computed for each stock regressing $\mathcal{W}_i \sim N^{\gamma}$: in average $\langle \gamma \rangle = 0.5$ as shown by the dashed black line. (Bottom right panel) Signature scatter plot (coloured by density of data) of the coefficients $\nu + \delta$ and γ respectively estimated conditioning to each stock.

over all the days and stocks. For each stock we regress $\mathcal{W}_i \sim N^{\gamma}$, $\sigma_{\mathrm{d}} \sim N^{\nu}$, $Q_i S_i \sim N^{\delta}$, and we obtain from the empirical distributions of the exponents in Figure 7.7 that their average values read $\langle \gamma \rangle = 0.5$, $\langle \nu \rangle = 0.25$ and $\langle \delta \rangle = 0.20$, thus $\langle \gamma \rangle \neq \langle \nu \rangle + \langle \delta \rangle$. However, by looking at the scatter plot of the estimated exponent γ as function of the sum $\nu + \delta$ computed separately for each stock (see bottom right panel in Figure 7.7) one observes a clear linear relation.

A possible and intuitive explanation of the non null measured correlation between σ_{d} and $Q_i S_i$ is that metaorders add up to volume, generate market impact and thus increase price volatility. In this way trading volume increases due to both an increase in the number of metaorders and in their sizes, and so does volatility from the increased market impact as discussed for example in [59]. Note that this reasoning is valid even if the metaorders only account for a certain percentage of the total daily market volume $V = \sum_{i=1}^{N} Q_i = \upsilon V_\mathrm{d}$ with υ adjusting for the partial view of the ANcerno sample in terms of volume, and for the non-*bet* traded by intermediaries: from our dataset we measure in average $\langle \upsilon \rangle \approx 5 \times 10^{-2}$.

7.C Statistics of exchanged risks and trading costs

Here we describe some statistics of the metaorders executed from the main investments funds and brokerage firms gathered by ANcerno. The empirical probability distribution of the number of metaorders N per asset, of the risk $\mathcal{W}_i$ exchanged by a metaorder and of the total daily traded risk $\mathcal{W}$ per asset are illustrated in Figure 7.8. It emerges that both the number of daily metaorders N and the risk measures typically vary over several orders of magnitude. In particular, as evident from the left panel in Figure 7.8, there is a significant number of metaorders active every day, since in average ~ 5 metaorders are executed per day for each asset. Furthermore, as shown in the right panel of Figure 7.8, both the single metaorder's risk $\mathcal{W}_i$ and the total daily exchanged risk $\mathcal{W}$ vary over almost eight decades. Note that these statistical properties are approximately independent from the time period and from the economical sector of the asset exchanged through metaorders.

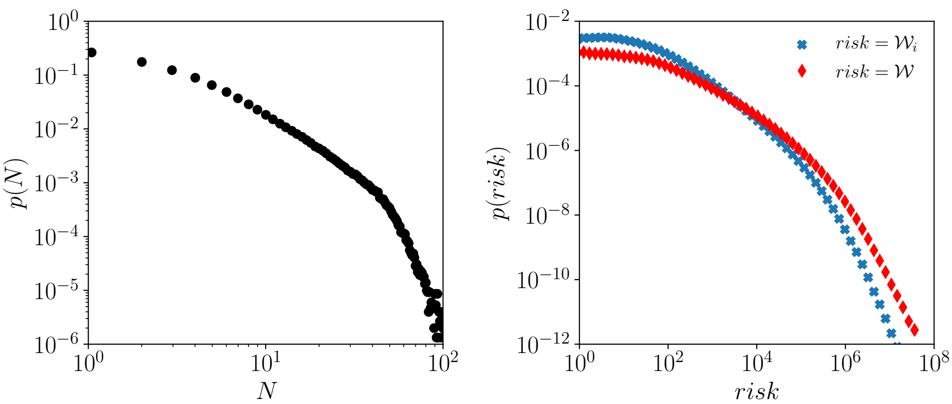

Figure 7.8: (Left panel) Empirical probability distribution of the daily number N of metaorders per asset: N is broadly distributed over two decades with an average close to 5. (Right panel) Empirical probability distributions of the exchanged risk per metaorder, i.e $\mathcal{W}_i = \sigma_\mathrm{d} Q_i S_i$, and of the total daily risk per day/assets, i.e $\mathcal{W} = \sum_{i=1}^{N} \mathcal{W}_i$.

In terms of the trading cost statistics we find that, for a single metaorder with unsigned volume Q, the spread cost $c_\mathrm{spd} = \mathcal{S} \times Q$ is dominant for small volumes, while the market impact cost $c_\mathrm{imp} = \sigma_\mathrm{d} \times SQ \times \sqrt{Q/V_\mathrm{d}}$ takes over for large volumes (see left panel of Figure 7.9). Furthermore, as shown in the right panel of Figure 7.9, the average daily market impact cost $\mathcal{C}_\mathrm{imp}$ accounts on average for $\approx 1/2$ of the total daily trading average cost $\mathcal{C} = \mathcal{C}_{\text{spd}} + \mathcal{C}_{\text{imp}}$, computed using $Y_{\text{spd}} = 3.5$ and $Y_{\text{imp}} = 1.5$ in Eq. (7.5).

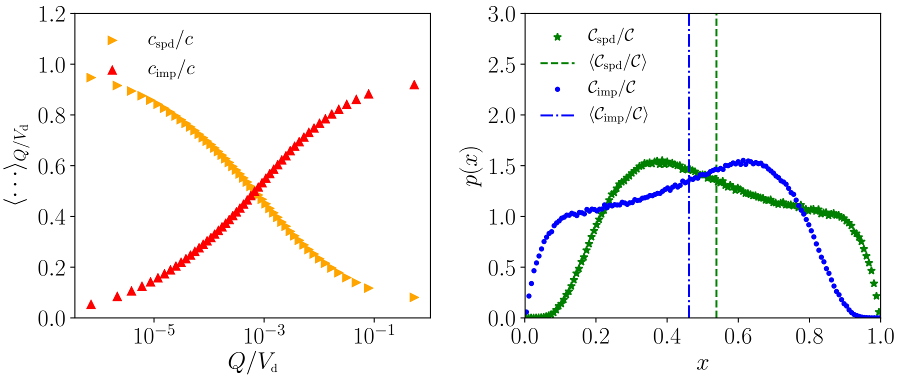

Figure 7.9: (Left panel) Averaged spread and market impact cost ratios given respectively by c_{spd}/c and c_{imp}/c - with $c_{\text{spd}} = \mathcal{S} \times Q$ (spread cost), $c_{\text{imp}} = \sigma_{\text{d}} \times SQ \times \sqrt{Q/V_{\text{d}}}$ (market impact cost) and $c = c_{\text{spd}} + c_{\text{imp}}$ (total cost per metaorder) - as function of the metaorder's order size Q/V_{d}: to note that for a metaorder with small (large) order size the spread (market impact) cost is dominant. (Right panel) Empirical distributions of the $\mathcal{C}_{\text{spd}}/\mathcal{C}$ and $\mathcal{C}_{\text{imp}}/\mathcal{C}$ ratios which give us an idea of the order of magnitude of the different contributions to the total daily average cost per metaorder $\mathcal{C} = \mathcal{C}_{\text{spd}} + \mathcal{C}_{\text{imp}}$ (computed from Eq. (7.5) fixing $Y_{\text{spd}} = 3.5$ and Y_{imp}=1.5): the dashed vertical lines represent the location of the mean values equal respectively to $\langle \mathcal{C}_{\text{spd}}/\mathcal{C} \rangle = 0.49$ and $\langle \mathcal{C}_{\text{imp}}/\mathcal{C} \rangle = 0.51$.

Conclusions

It is widely recognized that trading costs cause a significant drag on fund performance reducing the profit of a trading strategy. In particular, for large institutional investors these trading costs are mainly represented by the market impact - the mechanism through which trades move prices. One of the most surprising empirical findings is that the market impact of a so-called metaorder, i.e. a long sequence of orders executed incrementally in the same direction by the same investor, is approximately described by a square-root law. Given the not linear nature of this statistical law, prices can be affected by emergent effects related to the interaction between different investors following the same trading signal and using similar portfolio allocation strategies. For this reason the principal aim of this thesis is to investigate the market impact with a data-driven approach based on the ANcerno dataset which contains metaorders executed by an heterogeneous set of investors in the U.S. equity market.

In Chapter 4 we quantitatively test for the first time the dynamical theory of liquidity proposed in [69]. In perfect agreement with this theoretical framework, we find that the price change conditioned to the incremental execution of an order is characterized by a crossover in volume from a linear to a square-root regime. Taking into consideration the interaction of market participants having a wide spectrum of reaction time scales allows us to reproduce quantitatively and accurately the observed crossover from a linear to a square-root impact. Furthermore, we find that the square-root regime is independent from the execution duration of the metaorder, in agreement also with the findings discussed in Appendix A where we argue that market impact should not be misconstrued as volatility.

In Chapter 5 we shed light on an apparent paradox: How can a non-linear impact law survive in the presence of a large number of simultaneously executed metaorders? We answer this question introducing and investigating the co-impact, which is the term introduced to describe the market impact taking in consideration the crowding effects between metaorders simultaneously executed on the same asset. We find that the market chiefly reacts to the net order flow of ongoing metaorders, without individually distinguishing them. The joint co-impact of multiple contemporaneous metaorders depends on the total number of metaorders and their mutual sign correlation. Using a simple heuristic model calibrated on data, we reproduce very well the different regimes of the empirical market impact curves as a function of the volume fraction ϕ: square-root for large ϕ, linear for intermediate ϕ, and a finite intercept I_0 when $\phi \to 0$. The value of the intercept I_0 grows with the sign correlation coefficient. For an uncorrelated metaorder with the rest of the market, the impacts of other metaorders cancel out on average. On the contrary, any intercept of the impact law can be interpreted as a non-zero correlation with the rest of the market.

In Chapter 6 we investigate what happens to the price when the metaorder execution is completed. We find that price relaxation takes place as soon as the metaorder ends. While at the end of the same execution day it is on average $\approx 2/3$ of the peak impact, the decay continues the next days, following a power-law function at short time scales, and apparently converges to a non-zero asymptotic value at long time scales (~ 50 days) equal to $\approx 1/2$ of the impact at the end of the first day. Due to a significant, multiday correlation of the sign of executed metaorders, we find that a careful deconvolution of the observed impact must be performed to extract the estimated impact decay of an isolated metaorder.

In Chapter 7 we revisit the trading invariance hypothesis recently proposed by Kyle and Obizhaeva [51,52] identifying their bets with the ANcerno metaorders. The trading invariance hypothesis predicts that the quantity $I_{\text{ko}} := \mathcal{W}/N^{3/2}$, where $\mathcal{W}$ is the exchanged risk (volatility x volume x price) and N is the number of bets, is invariant. We find that the 3/2- scaling between $\mathcal{W}$ and N works well and it is robust against changes of year, market capitalisation and economic sector. However, our analysis clearly shows that I_{ko} is not invariant and that it is highly correlated with the total trading cost (spread and market impact) of the metaorder. We then propose new invariants defined as a ratio of I_{ko} and costs finding a large decrease in variance reconducible to the scaling of the spread with the volatility per transaction, the near invariance of the distribution of metaorder size and of the number fractions of metaorders across stocks. Based on the observed empirical results we argue that the trading invariance is a consequence of the validity of the square-root law for market impact as well as to the proportionality between spread and volatility.

In conclusion, this thesis presents a selection of studies related to the market impact of metaorders executed in the U.S. equity market. We would like to point out that the reason why large orders have smaller than expected (concave) market impact is far from being trivial. A lot of efforts in the literature are devoted to understand to what extent the size of the metaorder allows to predict the price change that they are going to induce, and yet there is no overall consensus with respect to the precise reasons leading to this behavior. From our side, we push the boundaries of the problem in several directions and with different approaches. At this point it would be interesting to investigate the validity of our results in other markets such as fixed income, currency or future markets. Finally, we want to stress that given the importance of the discussed topics, this manuscript is of interest to academicians that investigate price formation mechanism, to practitioners trying to monitor and reduce their trading costs, and also to regulators that seek to improve the stability of markets.

Appendix A

Impact is not just volatility

The notion of market impact is subtle and sometimes misinterpreted. Here we argue that impact should not be misconstrued as volatility. In particular, the so-called square-root impact law, which states that impact grows as the square-root of traded volume, has nothing to do with price diffusion, i.e. that typical price changes grow as the square-root of time.

The importance of market impact is strictly related to the one of the Transaction Cost Analysis which has become a very relevant issue in the Asset Management industry. Transaction costs account for a substantial part of the profits or losses of any investment strategy. When executing an order to buy or to sell, investors and trading firms have to worry about several sources of costs. Some costs are easy to identify and quantify, like market fees or spread costs. Much more subtle, but dominant for large portfolios, are impact costs. Intuitively, market impact describes the fact that, on average, buy orders tend to push the price up and sell orders tend to drag the price down. All the subtlety, however, lies in the words *on average*. Clearly, while our putative investor is executing his/her buy order, many things can happen: other investors may simultaneously buy or sell, market-makers/high frequency traders may unload their inventories, or some news may become available, pushing the price up or down. While some of these events may be directly related to his/her buy order, most of them result in a price move unbeknownst to our investor.

For a large enough number of executed orders, these random price moves average to zero. But for any given order executed within a time T, the price will randomly move up or down by an amount $\sim \sigma_{\mathrm{d}}\sqrt{T}$, where σ_{d} is the volatility. The impact of an order, on the other hand, is the part of the price move that survives upon averaging. Not surprisingly, this impact is much smaller than $\sigma_{\mathrm{d}}\sqrt{T}$ for small order sizes – see below for more precise statements. This definition of impact should be further refined to distinguish between the reactional (or mechanical) impact, that would exist even for trades without any information content, and the prediction related impact that reveals the information content of the trade, see [34]. The latter component is usually very small for medium to long term investors, since information (if any) is supposed to affect investment time scales much longer than the execution time T (typically several weeks or months compared to $T \sim$ a few days at most[1].

One interesting question concerns the dependence of the reactional impact $I(Q,T)$ on the size Q and duration T of the executed order. A now commonly accepted result is the so-called square-root law (see e.g. [2, 5, 9, 10, 12, 13, 37, 38, 50, 82]) which states that in normal trading conditions

$$I(Q,T) \approx Y \sigma_{\mathrm{d}} \sqrt{\frac{Q}{V_{\mathrm{d}}}}, \tag{A.1}$$

where Y is a constant of order unity, σ_{d} and V_{d} are, respectively, the daily volatility and daily volume corresponding to the traded asset. In fact, a more accurate description was proposed theoretically in [7, 69], where this square-root dependence becomes linear for small Q; more precisely:

$$I(Q,T) = \sigma_{\mathrm{d}} \sqrt{\frac{Q}{V_{\mathrm{d}}}}\, \mathcal{F}(\psi), \qquad \psi := \frac{Q}{V_{\mathrm{d}} T}, \tag{A.2}$$

where T is expressed in days such that $V_{\mathrm{d}}T$ is the total volume executed during T. The scaling function $\mathcal{F}(\psi)$ is monotonic and behaves as $\sqrt{\psi}$ for $\psi \to 0$ and as a constant Y for $\psi \to \infty$. Therefore, $I(Q,T)$ is linear in Q for small Q at fixed T, and crosses over to a square-root for large Q [2]. This prediction appears to describe empirical data surprisingly well as discussed in Chapter 4.

Such a functional form for the reactional impact has two immediate consequences. One is that for $Q \ll V_{\mathrm{d}}T$ (i.e. when the volume of the whole order Q is much smaller than the market volume during time T), reactional impact is much smaller than the typical price moves: $I(Q,T) \ll \sigma_{\mathrm{d}}\sqrt{T}$. In other words, the average impact (and therefore the associated impact cost) incurred by our investor is very small compared to the uncertainty on the price move during execution. Large trading firms with active investment strategies will however be mostly sensitive to the former (as the latter averages to zero), whereas once-off investors will want to minimise uncertainty, for fear of an adverse price move while their order is executed. The latter concern is at the heart of the famous Almgren-Chriss formalism [26].

The second, somewhat surprising consequence is that in the square-root regime, reactional impact depends only weakly on the execution time T – whereas of course the typical price changes increase as $\sqrt{T}$. Now, this $\sqrt{T}$ dependence was recently argued to be the mechanism at the heart of the square-root impact law [39]. The argument, in a nutshell, is that typical investors are essentially sensitive to price uncertainty, so perceived costs behave as $\sigma_{\mathrm{d}}\sqrt{T}$. But if their order of size Q is executed at a constant rate m_0, the time needed to complete the order is $T = Q/m_0$. Hence apparent impact behaves as $\sigma_{\mathrm{d}}\sqrt{Q/m_0}$, i.e. a square-root dependence on Q. We believe that this argument is

[1]Most trades in the ANcerno database discussed below appear to belong to medium or long term investors.

[2]Note that, as discussed in [34], this behaviour is not expected to hold in extreme trading conditions, for example when ϕ is large and T is small. If latent liquidity has no time to reveal itself in the order book, convex impact or even runaway situations can ensue – see for example [40].

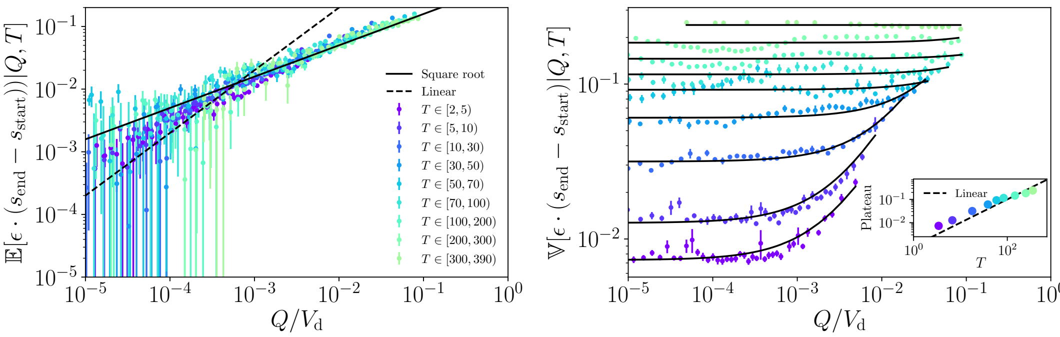

Figure A.1: Market impact curves $I(Q,T) = \mathbb{E}[\epsilon \cdot (s_{\text{end}} - s_{\text{start}})|Q,T]$ (left panel) and price uncertainty measured by $\mathbb{V}[\epsilon \cdot (s_{\text{end}} - s_{\text{start}})|Q,T]$ (right panel) as a function of the relative metaorder size Q/V_{d} for different buckets of order duration T: for both the panels and with an abuse of notation the log-prices $(s_{\text{start}}, s_{\text{end}})$ are rescaled by the daily volatility σ_{d} per stock while the order size is given by the ratio between the number of shares Q and the daily total market volume V_{d} per stock. As shown in the inset of the right panel, the plateau value increases linearly with the duration T, as expected for a random walk.

very misleading, and fails at explaining why reactional impact (i.e. the average of price moves, and not its root-mean-square) behaves as $\sqrt{Q}$ [3].

The difference between these two quantities is shown in Figure A.1, based on the ANcerno database (for full details see Chapter 3). In the left panel, we show the average log-price difference between the start and the end of the order execution, conditioned to the order size Q, and for different order durations T (different colors). One clearly sees the crossover between a T-dependent, steeper than square-root regime for small Q/V_{d} and a T-independent, square-root regime for larger Q/V_{d}, as already reported in [10, 35, 41, 42]. Note that over 80% of the empirical data lies in the square-root regime.

In the right panel, we show the variance of price differences, again conditioned to the order fraction and for different order durations. For small Q, this variance is nearly independent of the order size Q but linearly increases with T as expected (see inset). For larger Q, the variance acquires some dependence on Q/V_{d}, specially for small T. In order to rationalize these findings, let us postulate that the log-price change Δs between the start and the end of the execution of an order of duration T is given by the sum of

[3] More generally, the average overhead in costs due to impact $\mathcal{C}$ can be directly related to $I(Q,T)$ via the relation $\mathcal{C} = \int_0^T \mathrm{d}t\, \dot{Q}(t) I(Q(t), t)$, where $Q(t)$ is the quantity executed by the investor at each time $t \leq T$. On the other hand, the idiosyncratic price moves $\sim \sigma_{\text{d}} \sqrt{Q/m_0}$ only relate to execution risk, and are not directly linked to the average cost $\mathcal{C}$ paid by investors, for they average out after a sufficiently large number of investment decisions are executed.

an impact contribution and a volatility contribution, i.e.

$$\Delta s = s_{\rm end} - s_{\rm start} = \epsilon \cdot I(Q,T) \times (1 + \alpha\,\omega) + \sigma_{\rm d}\sqrt{T}\xi, \tag{A.3}$$

where $s_\circ = \log S_\circ$ with $\circ = \text{start}, \text{end}$ [4], $\epsilon = \pm 1$ depending on the sign of the order (buy or sell), α is a certain fitting parameter, ω, ξ are two independent random variables with zero mean and unit variance and T is, as above, measured as a fraction of the trading day. From this ansatz, it follows that

$$\mathbb{E}[\epsilon \cdot \Delta s | Q, T] = I(Q,T), \tag{A.4}$$

as it should be of course, and

$$\mathbb{V}[\epsilon \cdot \Delta s | Q, T] = \sigma_{\rm d}^2 T \left(1 + \alpha^2 \psi\, \mathcal{F}^2(\psi)\right). \tag{A.5}$$

This prediction is plotted in the right panel of Figure A.1, with α as the only fitting parameter ($\alpha \approx 10^{-1}$). Taken together, our ansatz and the two panels of Figure A.1 confirm that:

1. The square-root impact law for $\mathbb{E}[\epsilon \cdot \Delta s | Q, T]$ is completely unrelated to the scaling of the volatility as $\sqrt{T}$.
2. The square-root impact law and its fluctuations (parameterized by α) allows one to understand the systematic increase of volatility in the presence of a locally large order (large ψ).
3. As expected, price uncertainty (measured by $\mathbb{V}[\epsilon \cdot \Delta s | Q, T]$) largely exceeds the average reactional impact contribution: compare the square of the y-axis of the left panel with the y-axis of the right panel. This means, as emphasized in [39], that impact has a poor explanatory power compared to volatility. If an asset manager represents 1% of the market volume, he or she contributes to 5% of the volatility, which gives a coefficient of determination r^2 for the impact term of $\approx 2.5 \times 10^{-3}$.

In conclusion, we want to point out in this appendix some basic facts about market impact that are sometimes misinterpreted, with Ref. [39] as a case in point. We argue that impact should not be misconstrued as volatility and in particular, the so-called square-root impact law, which states that impact grows as the square-root of traded volume, has nothing to do with price diffusion, i.e. that typical price changes grow as the square-root of time. We rationalise empirical findings on impact and volatility by introducing a simple scaling argument which is in agreement with data. It follows that, even when execution risk is relevant for some investors (and at the core of the Almgren-Chriss formalism [26]), price variance should certainly not be misconstrued as price impact.

[4] Note that we are not rescaling the log-prices ($s_{\rm start}$, $s_{\rm end}$) by the daily volatility $\sigma_{\rm d}$ as done in the chapters of this thesis. This allows to put in evidence the explicit dependence on the daily volatility $\sigma_{\rm d}$ of the log-price change Δs as shown in Eq. (A.3).

Bibliography

[1] Kyle, A. S. (1985). continuous auctions and insider trading. Econometrica: Journal of the Econometric Society, 1315-1335.

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

[3] Gatheral, J. (2010). No-dynamic-arbitrage and market impact. Quantitative Finance, 10(7), 749-759.

[4] Almgren, R., Thum, C., Hauptmann, E., & Li, H. (2005). Direct estimation of equity market impact. Risk, 18(7), 58-62.

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

[6] Mastromatteo, I., Tóth, B., & Bouchaud, J. P. (2014). Anomalous impact in reaction-diffusion financial models. Physical Review Letters, 113(26), 268701.

[7] Donier, J., Bonart, J., Mastromatteo, I., & Bouchaud, J. P. (2015). A fully consistent, minimal model for non-linear market impact. Quantitative Finance, 15(7), 1109-1121.

[8] 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.

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

[10] Zarinelli, E., Treccani, M., Farmer, J. D., & Lillo, F. (2015). Beyond the square root: Evidence for logarithmic dependence of market impact on size and participation rate. Market Microstructure and Liquidity, 1(02), 1550004.

[11] Gomes, C., & 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.

[12] Brokmann, X., Serie, E., Kockelkoren, J., & Bouchaud, J. P. (2015). Slow decay of impact in equity markets. Market Microstructure and Liquidity, 1(02), 1550007.

[13] Tóth, B., Eisler, Z., & Bouchaud, J. P. (2016). The square-root impact law also holds for option markets. Wilmott, 2016(85), 70-73.

[14] Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. Nature, 423(6937), 267.

[15] Farmer, J. D., Gerig, A., Lillo, F., & Waelbroeck, H. (2013). How efficiency shapes market impact. Quantitative Finance, 13(11), 1743-1758.

[16] Jaisson, T. (2015). Market impact as anticipation of the order flow imbalance. Quantitative Finance, 15(7), 1123-1135.

[17] Bak, P., Paczuski, M., & Shubik, M. (1997). Price variations in a stock market with many agents. Physica A: Statistical Mechanics and its Applications, 246(3-4), 430-453.

[18] Bouchaud, J. P., Gefen, Y., Potters, M., & Wyart, M. (2004). Fluctuations and response in financial markets: the subtle nature of random price changes. Quantitative Finance, 4(2), 176-190.

[19] Bouchaud, J. P., Farmer, J. D., & Lillo, F. (2009). How markets slowly digest changes in supply and demand. In Handbook of financial markets: dynamics and evolution (pp. 57-160). North-Holland.

[20] Shiller, R. J. (1980). Do stock prices move too much to be justified by subsequent changes in dividends? National Bureau of Economic Research Cambridge, USA.

[21] Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2006). Institutional investors and stock market volatility. The Quarterly Journal of Economics, 121(2), 461-504.

[22] Lillo, F., Mike, S., & Farmer, J. D. (2005). Theory for long memory in supply and demand. Physical Review E, 71(6), 066122.

[23] Tóth, B., Palit, I., Lillo, F., & Farmer, J. D. (2015). Why is equity order flow so persistent? Journal of Economic Dynamics and Control, 51, 218-239.

[24] Beran, J. (2017). Statistics for long-memory processes. Routledge.

[25] Huberman, G., & Stanzl, W. (2004). Price manipulation and quasi-arbitrage. Econometrica, 72(4), 1247-1275.

[26] Almgren, R., & Chriss, N. (2001). Optimal execution of portfolio transactions. Journal of Risk, 3, 5-40.

[27] Bacry, E., & Muzy, J. F. (2014). Hawkes model for price and trades high-frequency dynamics. Quantitative Finance, 14(7), 1147-1166.

[28] O'hara, M. (1997). Market microstructure theory. Wiley.

[29] Harris, L. (2003). Trading and exchanges: Market microstructure for practitioners. OUP USA.

[30] Hasbrouck, J. (2007). Empirical market microstructure: The institutions, economics, and econometrics of securities trading. Oxford University Press.

[31] Vishwanath, R., & Krishnamurti, C. (2009). Investment management: A modern guide to security analysis and stock selection. Springer Science & Business Media.

[32] Abergel, F., Bouchaud, J. P., Foucault, T., Lehalle, C. A., & Rosenbaum, M. (2012). Market microstructure: confronting many viewpoints. John Wiley & Sons.

[33] Bouchaud, J. P. (2005). The subtle nature of financial random walks. Chaos: An Interdisciplinary Journal of Nonlinear Science, 15(2), 026104.

[34] Bouchaud, J. P., Bonart, J., Donier, J., & Gould, M. (2018). Trades, quotes and prices: financial markets under the microscope. Cambridge University Press.

[35] Bucci, F., Benzaquen, M., Lillo, F., & Bouchaud, J. P. (2019). Crossover from linear to square-Root market impact. Physical Review Letters, 122(10), 108302.

[36] Torre, N. (1997). BARRA market Impact model handbook. BARRA Inc., Berkeley.

[37] Donier, J., & Bonart, J. (2015). A million metaorder analysis of market impact on the Bitcoin. Market Microstructure and Liquidity, 1(02), 1550008.

[38] Frazzini, A., Israel, R., & Moskowitz, T. J. (2018). Trading costs. Available at SSRN: https://ssrn.com/abstract=3229719.

[39] Capponi, F., Cont, R., & Sani, A. (2019). Trade Duration, Volatility and Market Impact. Available at SSRN: https://ssrn.com/abstract=3351736.

[40] Dall'Amico, L., Fosset, A., Bouchaud, J. P., & Benzaquen, M. (2019). How does latent liquidity get revealed in the limit order book? Journal of Statistical Mechanics: Theory and Experiment, 2019(1), 013404.

[41] Bucci, F., Mastromatteo, I., Eisler, Z., Lillo, F., Bouchaud, J. P., & Lehalle, C. A. (2018). Co-impact: Crowding effects in institutional trading activity. Quantitative Finance, 1-13.

[42] Bucci, F., Benzaquen, M., Lillo, F., & Bouchaud, J. P. (2019). Slow decay of impact in equity markets: Insights from the ANcerno database. Market Microstructure and Liquidity.

[43] Puckett, A., & Yan, X. S. (2008). Short-term institutional herding and its impact on stock prices. Puckett, Available at SSRN: https://ssrn.com/abstract=972254.

[44] Goldstein, M. A., Irvine, P., Kandel, E., & Wiener, Z. (2009). Brokerage commissions and institutional trading patterns. The Review of Financial Studies, 22(12), 5175-5212.

[45] Chemmanur, T. J., He, S., & Hu, G. (2009). The role of institutional investors in seasoned equity offerings. Journal of Financial Economics, 94(3), 384-411.

[46] Torre, N., & Ferrari, M. J. (1998). The market impact model. Horizons, The Barra Newsletter, 165.

[47] Jame, R. (2010). Organizational structure and fund performance: Pension funds vs. mutual funds. Mutual Funds. Available at SSRN: https://ssrn.com/abstract=1540533.

[48] Puckett, A., & Yan, X. (2011). The interim trading skills of institutional investors. The Journal of Finance, 66(2), 601-633.

[49] Busse, J. A., Green, T. C., & Jegadeesh, N. (2012). Buy-side trades and sell-side recommendations: Interactions and information content. Journal of Financial Markets, 15(2), 207-232.

[50] Robert, E., Robert, F., & Jeffrey, R. (2012). Measuring and modeling execution cost and risk. The Journal of Portfolio Management, 38(2), 14-28.

[51] Kyle, A. S., & Obizhaeva, A. A. (2017). Dimensional analysis, leverage neutrality, and market microstructure invariance. Available at SSRN: https://ssrn.com/abstract=2785559.

[52] Kyle, A. S., & Obizhaeva, A. A. (2016). Market microstructure invariance: Empirical hypotheses. Econometrica, 84(4), 1345-1404.

[53] Kyle, A. S., Obizhaeva, A. A., & Tuzun, T. (2016). Microstructure invariance in US stock market trades. Available at SSRN: https://ssrn.com/abstract=1107875, FEDS Working Paper.

[54] Bae, K. H., Kyle, A. S., Lee, E. J., & Obizhaeva, A. A. (2016). Invariance of buy-sell switching points. Robert H. Smith School Research Paper No. RHS, 2730770.

[55] Bowe, M., Rizopoulos, E., & Zhang, S. S. Market microstructure invariance in the FTSE 100. Available at https://www.fmaconferences.org/Norway/Papers/MMIPaper.pdf.

[56] Andersen, T. G., Bondarenko, O., Kyle, A. S., & Obizhaeva, A. A. (2018). Intraday trading invariance in the E-mini S&P 500 futures market. Available at SSRN: https://ssrn.com/abstract=2693810.

[57] Benzaquen, M., Donier, J., & Bouchaud, J. P. (2016). Unravelling the trading invariance hypothesis. Market Microstructure and Liquidity, 2(03n04), 1650009.

[58] Pohl, M., Ristig, A., Schachermayer, W., & Tangpi, L. (2018). Theoretical and empirical analysis of trading activity. Mathematical Programming, 1-30.

[59] Jones, C. M., Kaul, G., & Lipson, M. L. (1994). Transactions, volume, and volatility. The Review of Financial Studies, 7(4), 631-651.

[60] Rogers, L. C. G., & Satchell, S. E. (1991). Estimating variance from high, low and closing prices. The Annals of Applied Probability, 504-512.

[61] Wyart, M., Bouchaud, J. P., Kockelkoren, J., Potters, M., & Vettorazzo, M. (2008). Relation between bid–ask spread, impact and volatility in order-driven markets. Quantitative Finance, 8(1), 41-57.

[62] Miller, M., & Modigliani, F. (1958). The cost of capital. Corporate Finance and the Theory of Investment. American Economic Review, 48, 261-297.

[63] Madhavan, A., Richardson, M., & Roomans, M. (1997). Why do security prices change? A transaction-level analysis of NYSE stocks. The Review of Financial Studies, 10(4), 1035-1064.

[64] Mandelbrot, B. B., & Stewart, I. (1998). Fractals and scaling in finance. Nature, 391(6669), 758-758.

[65] Gabaix, X. (2009). Power laws in economics and finance. Annual Review of Economics, 1(1), 255-294.

[66] Pohl, M., Ristig, A., Schachermayer, W., & Tangpi, L. (2017). The amazing power of dimensional analysis: Quantifying market impact. Market Microstructure and Liquidity, 3(03n04), 1850004.

[67] Buckingham, E. (1914). On physically similar systems; illustrations of the use of dimensional equations. Physical Review 4, 345.

[68] Kyle, A. S., & Obizhaeva, A. A. (2019). Market microstructure invariance: A dynamic equilibrium model. Available at SSRN: https://ssrn.com/abstract=2749531.

[69] Benzaquen, M., & Bouchaud, J. P. (2018). Market impact with multi-timescale liquidity. Quantitative Finance, 18(11), 1781-1790.

[70] Busse, J. A., Chordia, T., Jiang, L., & Tang, Y. (2016). Mutual fund transaction costs. Unpublished Working Paper, Emory University.

[71] Chakrabarty, B., Moulton, P. C., & Trzcinka, C. (2017). The performance of short-term institutional trades. Journal of Financial and Quantitative Analysis, 52(4), 1403-1428.

[72] Goldstein, M. A., Irvine, P., & Puckett, A. (2011). Purchasing IPOs with commissions. Journal of Financial and Quantitative Analysis, 46(5), 1193-1225.

[73] Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. Econometrica: Journal of the Econometric Society, 1029-1054.

[74] Tang, L. H., & Tian, G. S. (1999). Reaction–diffusion–branching models of stock price fluctuations. Physica A: Statistical Mechanics and its Applications, 264(3-4), 543-550.

[75] Smith, E., Farmer, J. D., Gillemot, L. S., & Krishnamurthy, S. (2003). Statistical theory of the continuous double auction. Quantitative Finance, 3(6), 481-514.

[76] Bouchaud, J. P., Kockelkoren, J., & Potters, M. (2006). Random walks, liquidity molasses and critical response in financial markets. Quantitative Finance, 6(02), 115-123.

[77] Kyle, A. S., & Obizhaeva, A. A. (2016). Large bets and stock market crashes. Available at SSRN: https://ssrn.com/abstract=2023776.

[78] Bacry, E., Dayri, K., & Muzy, J. F. (2012). Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. The European Physical Journal B, 85(5), 157.

[79] Eisler, Z., Kertesz, J., Lillo, F., & Mantegna, R. N. (2009). Diffusive behavior and the modeling of characteristic times in limit order executions. Quantitative Finance, 9(5), 547-563.

[80] Lillo, F. (2007). Limit order placement as an utility maximization problem and the origin of power law distribution of limit order prices. The European Physical Journal B, 55(4), 453-459.

[81] Jame, R. (2017). Liquidity provision and the cross section of hedge fund returns. Management Science, 64(7), 3288-3312.

[82] Mastromatteo, I., Tóth, B., & Bouchaud, J. P. (2014). Agent-based models for latent liquidity and concave price impact. Physical Review E, 89(4), 042805.

[83] Tóth, B., Lillo, F., & Farmer, J. D. (2010). Segmentation algorithm for non-stationary compound Poisson processes. The European Physical Journal B, 78(2), 235-243.

[84] Wolfram, S. (1991). Mathematica: A system for doing mathematics by computer. Addison-Wesley Redwood City (Calif.).

[85] Wong, R. (2001). Asymptotic approximations of integrals (Vol. 34). SIAM.

[86] Lillo, F., & Farmer, J. D. (2004). The long memory of the efficient market. Studies in Nonlinear Dynamics & Econometrics, 8(3).

[87] Jusselin, P., & Rosenbaum, M. (2018). No-arbitrage implies power-law market impact and rough volatility. Available at SSRN: https://ssrn.com/abstract=3180582.

[88] Said, E., Ayed, A. B. H., Husson, A., & Abergel, F. (2017). Market Impact: A systematic study of limit orders. Market Microstructure and Liquidity, 3(03n04), 1850008.

[89] Donier, J., & Bonart, J. (2015). A million metaorder analysis of market impact on the Bitcoin. Market Microstructure and Liquidity, 1(02), 1550008.

[90] Sağlam, M., Moallemi, C. C., & Sotiropoulos, M. G. (2019). Short-term trading skill: An analysis of investor heterogeneity and execution quality. Journal of Financial Markets, 42, 1-28.

[91] Kanazawa, K., Sueshige, T., Takayasu, H., & Takayasu, M. (2018). Derivation of the Boltzmann equation for financial Brownian motion: Direct observation of the collective motion of high-frequency traders. Physical Review Letters, 120(13), 138301.

[92] Rak, R., Drozdz, S., Kwapien, J., & Oswiecimka, P. (2013). Stock returns versus trading volume: Is the correspondence more general? arXiv preprint arXiv:1310.7018.

[93] Patzelt, F., & Bouchaud, J. P. (2018). Universal scaling and nonlinearity of aggregate price impact in financial markets. Physical Review E, 97(1), 012304.

[94] Kubilius, K., Mishura, I. S., & Ralchenko, K. (2017). Parameter estimation in fractional diffusion models. Springer.

[95] Zhang, Y. C. (1999). Toward a theory of marginally efficient markets. Physica A: Statistical Mechanics and its Applications, 269(1), 30-44.

[96] Ponzi, A., Lillo, F., & Mantegna, R. N. (2009). Market reaction to a bid-ask spread change: A power-law relaxation dynamics. Physical Review E, 80(1), 016112.

[97] Gayduk, R., & Nadtochiy, S. (2018). Liquidity effects of trading frequency. Mathematical Finance, 28(3), 839-876.

[98] Guéant, O. (2017). Optimal market making. Applied Mathematical Finance, 24(2), 112-154.

[99] Bonart, J., & Gould, M. D. (2017). Latency and liquidity provision in a limit order book. Quantitative Finance, 17(10), 1601-1616.

[100] Alfonsi, A., & Acevedo, J. I. (2014). Optimal execution and price manipulations in time-varying limit order books. Applied Mathematical Finance, 21(3), 201-237.

[101] Obizhaeva, A. A., & Wang, J. (2013). Optimal trading strategy and supply/demand dynamics. Journal of Financial Markets, 16(1), 1-32.

[102] Chan, L. K., & Lakonishok, J. (1993). Institutional trades and intraday stock price behavior. Journal of Financial Economics, 33(2), 173-199.

[103] Farmer, J. D., Gillemot, L., Lillo, F., Mike, S., & Sen, A. (2004). What really causes large price changes? Quantitative Finance, 4(4), 383-397.

[104] Farmer, J. D., & Zamani, N. (2007). Mechanical vs. informational components of price impact. The European Physical Journal B, 55(2), 189-200.

[105] Hopman, C. (2007). Do supply and demand drive stock prices? Quantitative Finance, 7(1), 37-53.

[106] Weber, P., & Rosenow, B. (2006). Large stock price changes: Volume or liquidity? Quantitative Finance, 6(1), 7-14.

[107] Weber, P., & Rosenow, B. (2005). Order book approach to price impact. Quantitative Finance, 5, 357-364.

[108] Malkiel, B. G., & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383-417.

[109] Grossman, S. J., & Stiglitz, J. E. (1976). Information and competitive price systems. The American Economic Review, 246-253.

[110] Lillo, F., Farmer, J. D., & Mantegna, R. N. (2003). Econophysics: Master curve for price-impact function. Nature, 421(6919), 129.

[111] Potters, M., & Bouchaud, J. P. (2003). More statistical properties of order books and price impact. Physica A: Statistical Mechanics and its Applications, 324(1-2), 133-140.

[112] Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. Econometrica : Journal of the Econometric Society, 135–155.

[113] Tauchen, G. E., & Pitts, M. (1983). The price variability-volume relationship on speculative markets. Econometrica : Journal of the Econometric Society, pages 485–505.

[114] Engle, R. F. (2000). The econometrics of ultra-high-frequency data. Econometrica, 68(1) :1–22.

[115] Eisler, Z., & Kertész, J. (2006). Size matters, some stylized facts of the market revisited. European Journal of Physics., B51 :145–154.

[116] Zumbach, G. (2004). How the trading activity scales with the company sizes in the FTSE 100. Quantitative Finance, 4 :441.

[117] Curtis, W., Logan, J. D., & Parker, W. (1982). Dimensional analysis and the pi theorem. Linear Algebra and its Applications, 47:117-126.