# 12

# The Impact of Metaorders

*It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.*

(Richard P. Feynman)

In the previous chapter, we considered how the arrival of a single market order impacts the mid-price. However, as we noted in Section 10.5.1, most traders do not execute large trades via single market orders, but instead split up their trades into many small pieces. These pieces are executed incrementally, using market orders, limit orders, or both, over a period of several minutes to several days. As we saw in the last chapter, the chaining of market orders greatly affects their impact. Therefore, understanding the impact of a single market order is only the first step towards understanding the impact of trading more generally. To develop a more thorough understanding, we must also consider the impact of **metaorders** (defined in Section 10.4.3).

The empirical determination of metaorder impact is an important experiment whose results, when measured properly, are of great interest to academics, investors and market regulators alike. From a fundamental point of view, how does a metaorder of size $Q$ contribute to price formation? From the point of view of investors, what is the true cost of performing such a trade? How does it depend on market conditions, execution strategies, time horizons, and so on? From the point of view of regulators, can large metaorders destabilise markets? Is marked-to-market accounting wise when, as emphasised above, the market price is (at best) only meaningful for infinitesimal volumes?

Naively, it might seem intuitive that the impact of a metaorder should scale linearly in its total size $Q$. Indeed, as we will discuss in this chapter, many simple models of price impact predict precisely a linear behaviour. Perhaps surprisingly, empirical analysis reveals that in real markets, this scaling is not linear, but rather is approximately square-root. Throughout this chapter, we present this square-root law of impact and discuss several of its important consequences.

## 12.1 Metaorders and Child Orders

Assume that a trader decides to buy or sell some quantity $Q$ of a given asset. Ideally, the trader would like to buy or sell this whole quantity immediately, at the market price. However, as we discussed in Section 10.5.1, unless $Q$ is smaller than the volume available for immediate purchase or sale at the opposite-side best quote, then conducting the whole trade at the bid- or ask-price is not possible. Therefore, there is no such thing as a "market price" for a trade, because such a price can only be guaranteed to make sense for a very small $Q$. For the volumes typically executed by large financial institutions, there is rarely enough liquidity in the whole LOB to match the required quantity $Q$ all at once. Therefore, traders must split their desired metaorder (i.e. the full quantity $Q$) into many smaller pieces, called **child orders**, which they submit gradually over a period of minutes, hours, days and even months.

It is common practice for traders to decompose their trading activity into two distinct stages:

(i) The *investment-decision stage,* during which the trader determines the sign $\varepsilon$ (i.e. buy/sell) and volume $Q$ of the metaorder, and the desired time horizon $T$ over which to execute it, usually based on some belief about the future price of the asset.
(ii) The *execution stage,* during which the trader conducts the relevant trades to obtain the required quantity at the best possible price within the time window $T$.

The execution stage is sometimes delegated to a broker, who seeks to achieve specified execution targets such as VWAP or TWAP (see Section 10.5.2) by performing incremental execution of the metaorder in small chunks. As we will discuss further in Chapter 19, this incremental execution seeks to make use of the gradual refilling of the LOB from previously undisclosed orders. In principle, this should considerably improve the price obtained for the metaorder execution when compared to submitting large market orders that penetrate deep into the LOB (and that could even destabilise the market).

Importantly, the direction, volume and time horizon of a metaorder are determined and fixed during the investment-decision stage. Any decision to stop, extend or revert a metaorder comes from a new investment decision. As we discuss later in this chapter, it is important to keep this in mind to avoid spurious conditioning effects when measuring the impact of a metaorder.

## 12.2 Measuring the Impact of a Metaorder

Consider a metaorder with volume $Q$ and sign $\varepsilon$. Let $N$ denote the number of constituent child orders of this metaorder. For $i = 1, 2, \ldots, N$, let $t_i$, $\upsilon_i$ and $p_i$ denote, respectively, the execution time, volume and execution price of the $i^{\text{th}}$ child order. By definition, it follows that

$$\sum_{i=1}^{N} \upsilon_i = Q.$$

Because each of the $N$ child orders are part of the same metaorder, it follows that the sign of each child order is also equal to $\varepsilon$.[1]

As we discussed (in the context of market order impact) in Section 11.2, the ideal experiment for measuring the impact of a metaorder would involve comparing two different versions of history: one in which a given metaorder arrived, and one in which it did not, but in which all else was equal. In reality, however, this is not possible, so we must instead make do with estimating impact by measuring quantities that are visible in empirical data.

### *12.2.1 The Ideal Data Set*

To gain a detailed understanding of the impact of a metaorder, one would ideally have access to some form of proprietary data or detailed broker data that lists:

- which child orders belong to which metaorders;
- the values of $t_i$, $p_i$, and $\upsilon_i$ for each child order; and
- whether each child order was executed via a limit order or a market order.

With access to this ideal data, it would be straightforward to construct a detailed execution profile of a metaorder by recording the values of $t_i$, $p_i$ and $\upsilon_i$ for each child order. In reality, however, it is rare to have access to such rich and detailed data, so it is often necessary to cope with less detailed data, and to impose additional assumptions about metaorder execution.

### *12.2.2 Less-Detailed Data*

Even in the absence of the full information described in Section 12.2.1, it is still possible to gain insight into the impact of a metaorder, given only the following information:

- the sign $\varepsilon$ and total quantity $Q$ of the metaorder;
- the time of the first trade $t_1$ and the corresponding mid-price price $m_1$;[2]
- the time of the last trade $t_N$ and the corresponding mid-price price $m_N$. The execution horizon $T$ is then given by $T = t_N - t_1$.

[1] We assume here that the investor does not send child orders to sell (respectively buy) when s/he wants to buy (respectively sell). This is reasonable as any round-trip is usually costly – see Section 19.5.

[2] In principle, $t_1$ is different from the time $t_0$ at which the investment decision is made. However, we neglect here the short-term predictability that would lead to a systematic price change between these two times.

In the absence of more detailed information about the execution profile of a metaorder, we will rely on an important assumption: that, given $t_1$ and $t_N$, the execution profile of the metaorder's profile is approximately linear, such that the cumulative volume $q(t)$ executed up to time $t$ is

$$q(t) \approx \frac{t - t_1}{T} Q, \qquad q(t_N) \equiv Q. \tag{12.1}$$

In Section 12.2.3, we list several interesting properties of impact that can be measured within this simple framework. However, it is important to stress that measuring metaorder impact still requires relatively detailed data that indicates which child orders belong to which metaorders. This information is not typically available in most publicly available data, which is anonymised and provides no explicit trader identifiers. Using such data only allows one to infer the aggregate impact as in Section 11.4. Identifying aggregate impact with metaorder impact is misleading, and in most cases leads to a substantial underestimation of metaorder impact.

### *12.2.3 Impact Path and Peak Impact*

Given metaorder data of the type described in Section 12.2.2, and making the assumptions described in that section, one can define several quantities of interest for characterising the impact of a metaorder of total volume $Q$ and horizon $T$:

(i) The average **impact path** is the mean price path between the beginning and the end of a metaorder (see Figure 12.1):

$$\mathfrak{I}^{\text{path}}(q, t - t_1 | Q, T) = \langle \varepsilon \cdot (m_t - m_1) \,|\, q; Q, T \rangle, \tag{12.2}$$

where $q$ is the quantity executed between $t_1$ and $t$ (which, by Equation (12.1), we assume to grow approximately linearly with $t$). We will see later that the conditioning on $Q$ and $T$ can in fact be removed if the execution profile is linear. We introduce the notation $\mathfrak{I}$ for the impact of a metaorder to distinguish it from the impact $\mathcal{I}$ of a single market order, which we discussed in Chapter 11.

(ii) The average **peak impact** of a metaorder of size $Q$ executed over a time horizon $T$ is:

$$\mathfrak{I}^{\text{peak}}(Q, T) := \langle \varepsilon \cdot (m_N - m_1) \,|\, Q \rangle. \tag{12.3}$$

As alluded to above, for a metaorder executed at a constant rate, it is reasonable that there cannot be any difference between the mechanical impact of the volume $q$ of a partially executed metaorder and a fully executed metaorder of volume $q$, since the rest of the market does not know whether the metaorder is continuing or not. Hence, neglecting any prediction contribution coming from short-term signals, we expect that

$$\mathfrak{I}^{\text{path}}(q, t - t_1 | Q, T) \approx \mathfrak{I}^{\text{peak}}(q, t - t_1), \qquad \text{for all } q \le Q, t \le T.$$

Empirical data confirms that this equality does indeed hold (see Section 12.5).

(iii) The **execution shortfall** $\mathscr{C}$ (also called execution cost or "slippage") is the average difference between the price paid for each subsequent child order and the decision price (which we assume to be the price paid for the first child order; see previous footnote). Neglecting the spread contribution, one writes

$$\mathscr{C}(Q, T) = \left\langle \sum_i \varepsilon \upsilon_i \cdot (m_i - m_1) \middle| Q \right\rangle. \tag{12.4}$$

This is the volume-weighted average premium paid by the trader executing the metaorder.

(iv) **Impact after-effects** describe the mean price path for $t > t_N$ (i.e. after the metaorder has been fully executed; see Figure 12.1). At any given $t > t_N$, the impact after-effect of a metaorder can be decomposed into a **transient component** $\mathfrak{I}^{\text{trans.}}(Q, t)$ and a **permanent component** $\mathfrak{I}^{\infty}(Q)$, such that, when $t > T$,

$$\mathfrak{I}^{\text{path}}(Q, t) = \mathfrak{I}^{\text{trans.}}(Q, t) + \mathfrak{I}^{\infty}(Q), \tag{12.5}$$

with

$$\mathfrak{I}^{\text{trans.}}(Q, t \to \infty) = 0; \qquad \mathfrak{I}^{\infty}(Q) = \lim_{t \to \infty} \mathfrak{I}^{\text{path}}(Q, t).$$

Note that the permanent component $\mathfrak{I}^{\infty}(Q)$ receives two types of contribution:

(a) One coming from the prediction signal at the origin of the metaorder, called the prediction impact in Section 11.2;

(b) The other coming from the possibly permanent reaction of the market to all trades, even uninformed. This contribution is nicely illustrated by the zero-intelligence Santa Fe model, for which indeed $\mathfrak{I}^{\infty}(Q) > 0$.

For non-linear execution profiles, one should expect the impact path to depend not only on $q$, but possibly on the whole execution schedule $\{q(t)\}$.

## 12.3 The Square-Root Law

At the heart of most empirical and theoretical studies of **metaorder impact** lies a very simple question: how does the impact of a metaorder depend on its size $Q$? Many models, including the famous Kyle model (see Chapter 15), predict this relationship to be linear. Although this answer may appear intuitive, there now exists an overwhelming body of empirical evidence that rules it out, in favour of a

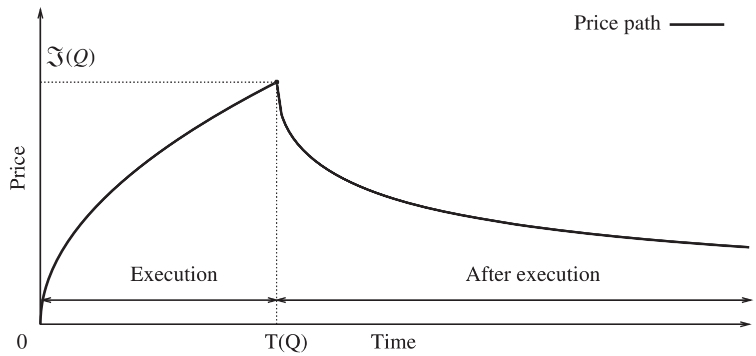


Figure 12.1. Average shape of the impact path. Over the course of its execution, a buy metaorder pushes the price up, until it reaches a peak impact. Upon completion, the buying pressure stops and the price reverts abruptly. Some impact is however still observable long after the metaorder execution is completed, and sometimes persists permanently. For real data, one should expect a large dispersion around the average impact path.

concave, and apparently square-root dependence on $Q$. In this section, we discuss the empirical basis of this so-called **square-root impact law**.

### *12.3.1 Empirical Evidence*

Since the early 1980s, a vast array of empirical studies[3] spanning both academia and industry have concluded that the impact of a metaorder scales approximately as the square-root of its size $Q$. This result is reported by studies of different markets (including equities, futures, FX, options, and even Bitcoin), during different epochs (including pre-2005, when liquidity was mostly provided by market-makers, and post-2005, when electronic markets were dominated by HFT), in different types of microstructure (including both small-tick stocks and large-tick stocks), and for market participants that use different underlying trading strategies (including fundamental, technical, and so on) and different execution styles (including using a mix of limit orders and market orders or using mainly market orders).

In all of these cases, the peak impact of a metaorder with volume $Q$ is well described by the relationship

$$\Im^{\text{peak}}(Q,T) \cong Y\sigma_T \left(\frac{Q}{V_T}\right)^{\delta}, \qquad (Q \ll V_T), \tag{12.6}$$

where $Y$ is a numerical coefficient of order 1 ($Y \cong 0.5$ for US stocks), $\delta$ is an exponent in the range 0.4–0.7, $\sigma_T$ is the contemporaneous volatility on the time horizon $T$, and $V_T$ is the contemporaneous volume traded over time $T$. Note that

[3] We list a wide range of such studies in Section 12.7.

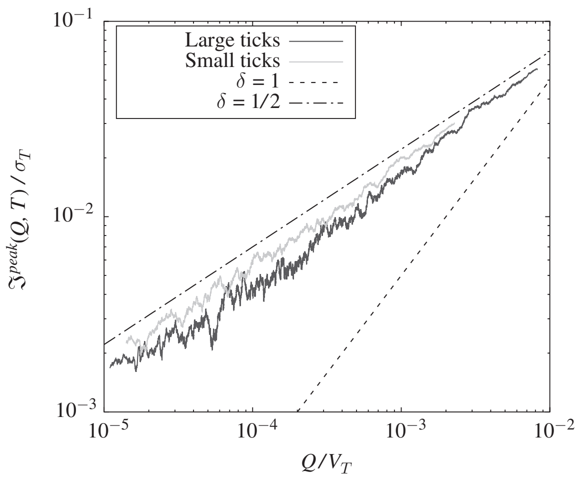


Figure 12.2. The impact of metaorders for Capital Fund Management proprietary trades on futures markets, during the period from June 2007 to December 2010 (see Tóth et al. (2011)). We show $\Im^{\text{peak}}(Q,T)/\sigma_T$ versus $Q/V_T$ on doubly logarithmic axes, where $\sigma_T$ and $V_T$ are the daily volatility and daily volume measured on the day the metaorder is executed. The black curve is for large-tick futures and the grey curve is for small-tick futures. For comparison, we also show a (dash-dotted) line of slope $\frac{1}{2}$ (corresponding to a square-root impact) and a (dotted) line of slope 1 (corresponding to linear impact).

Equation (12.6) is dimensionally consistent, in the sense that both the left-hand side and the right-hand side have the dimension [% of price].

To illustrate the relationship in Equation (12.6), Figure 12.2 shows $\Im^{\text{peak}}(Q,T)/\sigma_T$ vs. $Q/V_T$ for the data published in the paper by Tóth et al. in 2011,[4] which corresponds to nearly 500,000 metaorders executed between June 2007 and December 2010 on a variety of liquid futures contracts. The data suggests a value of $\delta \cong 0.5$ for small-tick contracts and $\delta \cong 0.6$ for large-tick contracts, for values of $Q/V_T$ ranging from about $10^{-5}$ to a few per cent. Several other empirical studies of different markets have drawn similar conclusions, with $\delta \cong 0.6$ for US and international stock markets, $\delta \cong 0.5$ for Bitcoin and $\delta \cong 0.4$ for volatility markets (see Section 12.7 for references). In all published studies, the exponent $\delta$ is around $\frac{1}{2}$, hence the name "square-root impact" that we will use henceforth.

### 12.3.2 A Very Surprising Law

The square-root law of metaorder impact is well established empirically, but there are several features that make it extremely surprising theoretically. In this section,

[4] 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.

we summarise some of these surprising features, and highlight some important lessons that the square-root law teaches us about financial markets.

The first surprising feature of Equation (12.6) is that metaorder impact does not scale linearly with $Q$ – or, said differently, that metaorder impact is not additive. Instead, one finds empirically that the second half of a metaorder impacts the price much less than the first half. This can only be the case if there is some kind of **liquidity memory time** $T_m$, such that the influence of past trades cannot be neglected for $T \ll T_m$ but vanishes for $T \gg T_m$, when all memory of past trades is lost. We will hypothesise in Chapter 18 that $T_m$ is in fact imprinted in the "latent" LOB that we alluded to above.

The second surprising feature of Equation (12.6) is that $Q$ appears not (as might be naively anticipated) as a fraction of the total market capitalisation $\mathcal{M}$ of the asset,[5] but instead as a fraction of the total volume $V_T$ traded during the execution time $T$. In modern equities markets, $\mathcal{M}$ is typically about 200 times larger than $V_T$ for $T = 1$ day. Therefore, the impact of a metaorder is much larger than if the $Q/V_T$ in Equation (12.6) was instead $Q/\mathcal{M}$. The square-root behaviour for $Q \ll V_T$ also substantially amplifies the impact of small metaorders: executing 1% of the daily volume moves the price (on average) by $\sqrt{1\%} = 10\%$ of its daily volatility. The main conclusion here is that even relatively small metaorders cause surprisingly large impact.

The third surprising feature of Equation (12.6) is that the time horizon $T$ does not appear explicitly. To consider this in more detail, let

$$\mathcal{L}_T := \sqrt{V_T}/\sigma_T \tag{12.7}$$

denote the **liquidity ratio**, which measures the capacity of the market to absorb incoming order flow. If prices are exactly diffusive, one has $\sigma_T = \sqrt{T}\sigma_1$, and if traded volume grows linearly with time, i.e. $V_T = TV_1$ (where "1" denotes, say, one trading day), then one finds

$$\mathcal{L}_T = \frac{\sqrt{TV_1}}{\sqrt{T}\sigma_1} = \frac{\sqrt{V_1}}{\sigma_1}.$$

Therefore, $\mathcal{L}_T$ is independent of $T$. If we rewrite Equation (12.6) to include $\mathcal{L}_T$, we arrive at (for $\delta = 0.5$)

$$\mathfrak{I}^{\text{peak}}(Q,T) = \frac{Y}{\mathcal{L}_T}\sqrt{Q}.$$

[5] The idea that trading 1% of a stock's total market capitalisation should move its price by about 1% was common lore in the 1980s, when impact was deemed totally irrelevant for quantities representing a few basis points of $\mathcal{M}$. Neglecting the potential impact of trades representing 100% of $V_T$, but (at the time) about 0.25% of the market capitalisation, is often cited as one of the reasons for the 1987 crash, when massive portfolio insurance trades created havoc in financial markets (see, e.g., Treynor, J. L. (1988). Portfolio insurance and market volatility. *Financial Analysts Journal*, 44(6), 71–73).

Hence, the square-root impact can be written as:

$$\mathfrak{I}^{\text{peak}}(Q,T) \cong Y\sigma_1 \sqrt{\frac{Q}{V_1}}, \qquad (Q \ll V_T), \tag{12.8}$$

which illustrates that the impact of a metaorder is only determined by $Q$, not the time $T$ that it took to be executed. A possible explanation, from an economic point of view, is that the market price has to adapt to a change of global supply/demand $\varepsilon Q$, independently of how this volume is actually executed. We provide a more detailed, Walrasian view of this non-trivial statement in Chapter 18.

### 12.3.3 Domain of Validity

As with all empirical laws, the square-root impact law only holds within a certain domain of parameters. In this section, we discuss this domain and highlight some of the limitations that one should expect on more general grounds.

First, as we noted in Section 12.3.2, the square-root law should only hold when the execution horizon $T$ is shorter than a certain memory time $T_m$, which is related to the time after which the underlying latent liquidity of the market has evolved significantly (see Chapter 18). The memory time $T_m$ is difficult to estimate empirically, but one may expect that it is of the order of days. Therefore, the impact of a metaorder of size $Q$ that is executed over a horizon of several weeks or months should become approximately proportional to $Q$, if the permanent component of impact is non-zero.

Second, the ratio $Q/V_T$ should be small, such that the metaorder constitutes only a small fraction of the total volume $V_T$ and such that the impact is small compared to the volatility. In the case where $Q/V_T$ is substantial, one enters a different regime, because the metaorder causes a large perturbation to the ordinary course of the market. In this large $Q/V_T$ regime, a metaorder with a sufficiently long execution horizon $T$ for the underlying latent liquidity to reveal itself is expected to have a completely different impact than a metaorder that is executed extremely quickly, with the risk of destabilising the market and possibly inducing a crash.

More formally, let $\tau_{\text{liq}}$ be the time needed for the LOB liquidity to refill, which is typically on the scale of seconds to minutes. The two above-mentioned regimes are:

- **Slow execution**, such that $\tau_{\text{liq}} \ll T \ll T_m$. For small $Q/V_T$, this is the square-root regime discussed in the last section. As $Q/V_T$ increases beyond (say) 10%, one may actually expect that the square-root behaviour becomes even more concave, because more and more sellers (respectively, buyers) eventually step in to buffer the increase (respectively, decrease) in price (see Equation (18.10) for a concrete example).[6]
- **Fast execution**, such that $T \lesssim \tau_{\text{liq}}$. In this case, impact becomes a convex (rather than concave) function of $Q$ for large $Q$. This is clear in the limit $T \to 0$, where the metaorder simply consumes the liquidity immediately available in the LOB. Since the mean volume profile in the LOB first increases then decreases with increasing distance from the mid-price (see Section 4.7), the immediate impact of a volume

[6] This seems consistent with empirical findings, see, e.g., Zarinelli et al. (2015). Beyond the square root: Evidence for logarithmic dependence of market impact on size and participation rate. *Market Microstructure and Liquidity*, 1(02), 1550004. One should however keep in mind the possible "Implementation Bias 1" discussed in the next section.

$Q$ will first be a concave, and then a convex, function of $Q$.[7] This is the reason that some financial markets implement circuit breakers, which aim to slow down trading and allow time for liquidity to replenish (see Section 22.2).

Finally, the regime where $Q$ is smaller than the typical volume at best $V_{\text{best}}$ is again expected to depart from a square-root behaviour and recover a linear shape. This is suggested both by some empirical results[8] and by the theoretical analysis of Chapters 18 and 19.

In summary, the square-root impact law holds (approximately) in the regime of intermediate execution horizons $T$ and intermediate volume fractions $Q/V_T$:

$$\tau_{\text{liq}} \ll T \ll T_m; \qquad \frac{V_{\text{best}}}{V_T} \ll \frac{Q}{V_T} \lesssim 0.1,$$

which is the regime usually adopted by investors in normal trading conditions. Different behaviours should be expected outside of these regimes, although data to probe these regions is scarce, and the corresponding conclusions currently remain unclear.

### *12.3.4 Possible Measurement Biases*

The peak impact of a metaorder $\Im^{\text{peak}}(Q,T)$ can be affected by several artefacts and biases. To obtain reproducible and understandable results, one should thus stick to a well-defined experimental protocol, and be aware of the following possible difficulties:

(i) **Prediction bias 1**: The larger the volume $Q$ of a metaorder, the more likely it is to originate from a stronger prediction signal. Therefore, a larger part of its impact may be due to short-term predictability, and not reveal any structural regularity of the reaction part of impact.

(ii) **Prediction bias 2**: Traders with strong short-term price-prediction signals may choose to execute their metaorders particularly quickly, to make the most of their signal before it becomes stale or more widely-known. Therefore, the strength of a prediction signal may itself influence the subsequent impact path (this is the prediction impact), in particular when the prediction horizon is comparable to the execution horizon. This bias is also likely to affect the long-term, permanent component $\Im^{\infty}(Q)$ of the impact (see Section 12.2).

(iii) **Synchronisation bias**: The impact of a metaorder can change according 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 correlated, or if they trade based on the same piece of news. (Note that this bias overlaps with the prediction biases discussed in (i) and (ii)).

(iv) **Implementation bias 1**: Throughout this chapter, we have assumed that both the volume $Q$ and execution horizon $T$ are fixed before a metaorder's execution begins. In reality, however, some traders may adjust these values during execution, by conditioning on the price path. In these cases, understanding metaorder impact is much more difficult. For example, when examining a buy metaorder that is only executed if the price goes down, and abandoned if the price goes up, impact will be negative. This implementation bias is expected to be stronger for large volumes, since the price is more likely to have

[7] Indeed, the immediate impact $\mathcal{I}$ is set by the condition:

$$\int_{m_1}^{m_1+\mathcal{I}} \mathrm{d}p'\, V_+(p') = Q,$$

where $V_+(p)$ is the density of available orders at price $p$ (see Section 3.1.7). Solving for $\mathcal{I}$ as a function of $Q$ leads to a convex shape at large $Q$ when $V_+(p)$ is a decreasing function of $p$.

[8] Zarinelli et al. (2015).

moved adversely in these cases. This conditioning may result in a systematic underestimation of $\mathfrak{I}^{\text{peak}}(Q,T)$ at large $Q$'s.

(v) **Implementation bias 2**: The impact path can become distorted for metaorders that are not executed at an approximately constant rate. For example, if execution is front-loaded, in the sense that most of the quantity $Q$ is executed close to the start time $t_1$, the impact at time $t_N$ will have had more time to decay, and may therefore be lower than for a metaorder executed at constant rate.

(vi) **Issuer bias**: Another bias may occur if a trader submits several dependent metaorders successively. If such metaorders are positively correlated and occur close in time to one another, the impact of the first metaorder will be different to the impact of the subsequent metaorders. Empirically, the impact of the first metaorder is somewhat greater than that of the second, and so on (much as for single market orders; see Section 11.3.4). Reasons for this will become clear in Chapter 19.

All these biases are pervasive. Fortunately, most of them can be avoided when one has access to proprietary trading data that contains full information about each child order, as described in Section 12.2.1. However, even less-ideal data sets still lead to similar conclusions on the dependence of the peak impact $\mathfrak{I}^{\text{peak}}(Q,T)$ on $Q$. Remarkably, many different studies (based on different assets, epochs, trading strategies and microstructure) all converge to a universal, square-root-like dependence of $\mathfrak{I}^{\text{peak}}(Q,T)$ on $Q$, and a weakly decreasing dependence on $T$.

### 12.3.5 Common Misconceptions

We now discuss some misconceptions and confusions about the impact of metaorders that exist in the literature.

- The impact of a metaorder of volume $Q$ is *not* equal to the aggregate impact of order imbalance $\Delta V$, which is linear, and not square-root-like, for small $Q$ (see Section 11.4).[9] Therefore, one cannot measure the impact of a metaorder without being able to ascribe its constituent trades to a given investor.
- The square-root impact law applies to slow metaorders composed of several individual trades, but *not* to the individual trades themselves. Universality, if it holds, can only result from some mesoscopic properties of supply and demand that are insensitive to the way markets are organised at the micro-scale. This would explain why the square-root law holds equally well in the pre-HFT era (say, before 2005) as today, why it is insensitive to the tick size, etc.
- Conversely, at the single-trade level, microstructure effects (such as tick size and market organisation) play a strong role. In particular, the impact of a single market order, $\mathcal{R}(\upsilon,1)$, does not behave like a square-root, although it behaves as a concave function of $\upsilon$. This concavity has no immediate relation with the concavity of the square-root impact for metaorders.
- A square-root law for the *mean impact* of a metaorder is not related to the fact that the average *squared* price difference on some time interval grows linearly with the total exchanged volume:

$$\mathbb{E}\left[(m_{t+T}-m_t)^2\middle|\sum_{n\in[t,t+T)}\upsilon_n = V\right] \propto V. \tag{12.9}$$

[9] In particular, the assumption that the aggregate order-flow imbalance $\Delta V$ in a given time window is mostly due to the presence of a single trader executing a metaorder of size $Q$, such that $\Delta V = Q+$ noise, is not warranted. For example, $\Delta V$ likely results from the superposition of *several* (parts of) overlapping metaorders with different sizes, signs and start times.

Equation (12.9) is a trivial consequence of the diffusive nature of prices when the total traded volume $V$ scales like $T$. Cursorily, this relationship reads as "price differences grow as the square-root of exchanged volumes", but it tells us nothing about the average *directional* price change in the presence of a directional metaorder.

The square-root law is a genuinely challenging empirical fact. Most models in the literature (such as the famous Kyle model; see Chapter 15) predict linear impact. We review some possible scenarios that could explain the concavity of metaorder impact in Section 12.6.

## 12.4 Impact Decay

At time $t_N$, the pressure exerted by the metaorder stops and the price reverts back (on average) towards its initial value. Empirical data again suggests a universal behaviour, at least shortly after $t_N$, when impact relaxes quite abruptly. The decay then slows down considerably, and impact appears to reach a plateau.

For times beyond $t_N$, empirical data becomes increasingly noisy, because the variance of price changes itself increases linearly with $t$. The issue of long-term impact is thus extremely difficult to settle, in particular because one expects impact to decay as a slow power-law of time to compensate the long memory of order signs (see Section 13.2 and Chapter 19). Therefore, although the transient, square-root impact $\mathfrak{I}^{\text{trans.}}(Q,t)$ is most likely universal, the permanent impact component $\mathfrak{I}^{\infty}(Q)$ is a combination of a reaction component (as would happen in the Santa Fe model) and a genuine prediction component (resulting from the fact that some agents do successfully predict the future price).[10] Perhaps tautologically, one expects $\mathfrak{I}^{\infty}(Q)$ to be higher for agents with a high predictive power than for noise trades (who trade without information, e.g. for risk or cash-flow purposes).

Since metaorders themselves tend to be autocorrelated in time (the same decision to buy or sell might still be valid on the next day, week or month), it is difficult to separate the mechanical contribution from the informational contribution, especially in view of the slow decay of impact. An apparent permanent impact could well be due to the continuing pressure of correlated metaorders. Again, measuring the amount of true information revealed by trades is very difficult (see Chapter 20).

## 12.5 Impact Path and Slippage Costs

How can we be sure that the square-root law is really universal, and is not a consequence of measurement biases (such as those listed in Section 12.3.4) or conditioning by traders, perhaps by some sophisticated optimisation program (see

[10] See references in Section 12.7.

Section 21.2.3)? Thankfully, high-quality proprietary data allows us to dismiss most of these concerns and to check that the square-root law is actually extremely robust. Indeed, one finds that the impact of the initial $\phi\%$ of a metaorder of total size $Q$ *also obeys the square-root law:*[11]

$$\mathfrak{I}^{\text{path}}(\phi Q, \phi T|Q,T) \cong \mathfrak{I}^{\text{peak}}(\phi Q) = \sqrt{\phi}\,\mathfrak{I}^{\text{peak}}(Q). \tag{12.10}$$

Hence, a partial metaorder consisting of the first $\phi\%$ of a metaorder of size $Q$ behaves as a *bona fide* metaorder of size $\phi Q$, and also obeys the square-root impact law. This implies that the square-root impact is not a mere consequence of how traders choose the size and/or the execution horizon of their metaorders as a function of their prediction signal (see Section 21.2.3 for a detailed discussion of this point).

Interestingly, Equation (12.10) allows one to relate the execution shortfall $\mathscr{C}$ (or impact-induced slippage) to the peak impact

$$\mathscr{C}(Q) = \left\langle \sum_i \varepsilon \upsilon_i \cdot (m_i - m_1) \middle| Q \right\rangle = \sum_i \upsilon_i \mathfrak{I}^{\text{path}}\left(\sum_{j<i} \upsilon_j\right). \tag{12.11}$$

Using Equation (12.10) and taking the continuous limit $\upsilon_i \to Q\mathrm{d}\phi$,

$$\mathscr{C}(Q) \approx Q \int_0^1 \mathrm{d}\phi\, \mathfrak{I}^{\text{path}}(\phi Q, \phi T|Q,T) \approx Q\mathfrak{I}^{\text{peak}}(Q) \int_0^1 \mathrm{d}\phi\, \sqrt{\phi} = \frac{2}{3} Q\mathfrak{I}^{\text{peak}}(Q).$$

A square-root impact therefore leads to a square-root execution shortfall. Importantly, the execution shortfall per unit volume corresponds to $\frac{2}{3}$ of the peak impact. If impact was linear, the execution shortfall per unit volume would only be $\frac{1}{2}$ of the peak impact.

## 12.6 Conclusion

The most important conclusion of this chapter is that, contrarily to intuition, the average impact of a metaorder does not scale linearly with its volume $Q$. Instead, impact obeys the square-root law from Equation (12.6). This behaviour is contrary to what is predicted by many models, such as the Kyle model (see Chapter 15), which predict linear behaviour. This is an interesting case where empirical data compelled the finance community to accept that reality was fundamentally different from mainstream theory.

Since the mid-nineties, several stories have been proposed to account for the square-root impact law. The first attempt, due to the Barra Group and Grinold

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

and Kahn (1999),[12] argues that the square-root behaviour is a consequence of market-markers being compensated for their inventory risk. The reasoning is as follows. Assume that a metaorder of volume $Q$ is absorbed by market-makers who will need to slowly offload their position later on. The amplitude of a potentially adverse move of the price during this unwinding phase is of the order of $\sigma_1 \sqrt{T_{\text{off}}}$, where $T_{\text{off}}$ is the time needed to offload an inventory of size $Q$. It is reasonable to assume that $T_{\text{off}}$ is proportional to $Q$ and inversely proportional to the trading rate of the market $V_1$ (see Section 12.3.2). If market-makers respond to the metaorder by moving the price in such a way that their profit is of the same order as the risk they take, then it would follow that $\Im \propto \sigma_1 \sqrt{Q/V_1}$, as found empirically. However, this story assumes no competition between market-makers. Indeed, inventory risk is diversifiable over time, and in the long run averages to zero. Charging an impact cost compensating for the inventory risk of each metaorder would lead to formidable profits and would necessarily attract competing liquidity providers, eventually leading to a Y-ratio much smaller than 1.

Another theory, proposed by Gabaix et al. (2003), ascribes the square-root impact law to the fact that the optimal execution horizon $T^*$ for informed metaorders of size $Q$ grows like $T^* \sim \sqrt{Q}$ (see Section 21.2.3 for a detailed derivation). This theory argues that since the price is expected to move linearly in the direction of the trade during $T^*$ (as information is revealed), the peak impact itself behaves as $\sqrt{Q}$. However, this scenario would imply that the impact path is linear in the executed quantity $q$, which is at odds with empirical data. As we noted in Section 12.5, the full impact path also behaves as a square-root of $q$ (at least when the execution schedule is flat).

Recently, Farmer et al. (2013) proposed yet another theory that is very reminiscent of the Glosten–Milgrom model, which argues that the size of the bid–ask spread is actually set competitively (see Section 16.1). The theory assumes that metaorders arrive sequentially, with a volume $Q$ distributed according to a power-law. Market-makers attempt to guess whether the metaorder will continue or stop at the next time step, and set the price such that it is a martingale and such that the average execution price compensates for the information contained in the metaorder (this condition is sometimes called "fair-pricing"). If the distribution of metaorder sizes behaves as $Q^{-5/2}$, these two conditions lead to a square-root impact law (see Sections 13.4.5 and 16.1.6 for more details). Although enticing, this theory has difficulty explaining why the square-root impact law appears to be much more universal than the distribution of the size of metaorders or of the autocorrelation of trade signs.[13] For example, the square-root law holds very precisely in Bitcoin markets, where the distribution of metaorder sizes behaves

[12] See also Section 12.7, in particular Zhang, Y.-C. (1999).
[13] On this last point, see Mastromatteo, I., Tóth, B., & Bouchaud, J. P. (2014). Agent-based models for latent liquidity and concave price impact. *Physical Review E*, 89(4), 042805.

as $Q^{-2}$, rather than $Q^{-5/2}$, and at a time where market-making was much less competitive.

Closer to the spirit of the present book, Tóth et al. (2011) proposed an alternative theory based on a dynamical description of supply and demand. This approach provides a natural statistical interpretation for the square-root law and its apparent universality. We provide a detailed overview of this idea in Chapters 18 and 19.

**Take-Home Messages**

(i) By studying the impact of metaorders, it is possible to quantify how market participants' actions impact prices at a mesoscopic scale.

(ii) According to the square-root law, the immediate price impact of a metaorder behaves as a square root of the order's volume.

(iii) The square-root law holds in a wide range of market scenarios, and can be observed empirically for both informed and uninformed metaorders.

(iv) When studying metaorder impact, it is important to avoid the many biases that could influence measurements.

(v) After a metaorder stops, the price reverts (on average) towards its pre-trade value. How much of the impact remains long after the execution (i.e. the permanent impact) depends on the predictive power or informational content of the metaorder.

## 12.7 Further Reading

Treynor, J. L. (1988). Portfolio insurance and market volatility. *Financial Analysts Journal*, 44(6), 71–73.

### *Empirical Evidence for a Square-Root Law*

Loeb, T. F. (1983). Trading cost: The critical link between investment information and results. *Financial Analysts Journal*, 39, 39–44.

Torre, N., & Ferrari, M. (1997). *Market impact model handbook*. BARRA Inc. Available at www.barra.com/newsletter/nl166/miminl166.asp.

Almgren, R., Thum, C., Hauptmann, E., & Li, H. (2005). Direct estimation of equity market impact. *Risk*, 18(5752), 10.

Kissel, R., & Malamut, R. (2006). Algorithmic decision-making framework. *Journal of Trading*, 1(1), 12–21.

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.

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.

Engle, R., Ferstenberg, R., & Russell, J. (2012). Measuring and modelling execution cost and risk. *The Journal of Portfolio Management*, 38(2), 14–28.

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

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

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.

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

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.

Tóth, B., Eisler, Z., & Bouchaud, J.-P. (2017). The short-term price impact of trades is universal. https://ssrn.com/abstract=2924029.

Several internal bank documents have also reported such a concave impact law, e.g.: Ferraris, A. (2008). *Market impact models*. Deutsche Bank internal document, http://dbquant.com/Presentations/Berlin200812.pdf.

## *Decay of Metaorder Impact*

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.

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

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.

see also: Moro, E. et al.; Zarinelli, E. et al. in the previous subsection.

## *Theories About the Square-Root Law*

Zhang, Y. C. (1999). Toward a theory of marginally efficient markets. *Physica A: Statistical Mechanics and Its Applications*, 269(1), 30–44.

Grinold, R. C., & Kahn, R. N. (2000). *Active portfolio management*. McGraw-Hill.

Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, 423(6937), 267–270.

Barato, A. C., Mastromatteo, I., Bardoscia, M., & Marsili, M. (2013). Impact of meta-order in the Minority Game. *Quantitative Finance*, 13(9), 1343–1352.

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

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

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.

Donier, J., & Bouchaud, J. P. (2016). From Walras auctioneer to continuous time double auctions: A general dynamic theory of supply and demand. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(12), 123406.

Pohl, M., Ristig, A., Schachermayer, W., & Tangpi, L. (2017). The amazing power of dimensional analysis: Quantifying market impact. arXiv preprint arXiv:1702.05434.

see also: K. Rodgers, https://mechanicalmarkets.wordpress.com/2016/08/15/price-impact-in-efficient-markets/.