Optimal Execution of Financial Transactions in Presence of Market Impact

Candidato:
Enzo Busseti

Relatore: Prof. Fabrizio Lillo
Relatore interno: Prof. Riccardo Mannella

Anno Accademico 2010/2011
## Contents

**Introduction**  

**Acronyms**  

1 Efficient market hypothesis, random walk, and market impact  
   1.1 Financial markets  
      1.1.1 Stocks  
      1.1.2 Stock markets  
      1.1.3 A mechanism for markets: the limit order book  
   1.2 Efficient Market Hypothesis  
   1.3 Prices, returns and random walk  
      1.3.1 Arithmetic random walk  
      1.3.2 Geometric random walk  
   1.4 Stylized facts of financial time series  
      1.4.1 Returns correlation  
      1.4.2 Fat tails  
      1.4.3 Volatility clustering  
   1.5 Why does the price move?  
      1.5.1 Effect of news  
      1.5.2 Market impact  
      1.5.3 Order flow  
      1.5.4 An apparent paradox  
   1.6 Consequences of correlated order flow  

2 Review on optimal execution  
   2.1 Motivations for optimal execution  
   2.2 Statement of the problem  
      2.2.1 Trading schedule  
      2.2.2 Execution costs  
   2.3 Optimal execution with permanent linear impact  
      2.3.1 Optimal solution  
      Minimum variance solution  
   2.4 Optimal execution with temporary impact and spread costs  
      2.4.1 Optimal solution  
   2.5 Risk aversion and efficient frontier of optimal execution
CONTENTS

Definition with Lagrange multiplier ........................................... 35

2.5.1 Optimal solutions for different parameters of risk aversion .......... 37

2.6 Optimal execution with exponentially decaying impact ................. 38

2.6.1 Optimal solution .................................................................. 39

3 Propagator model for market impact .......................................... 41

3.1 Motivations for the market impact propagator model ..................... 41

3.2 Definition of the propagator model ........................................ 42

3.2.1 Impact function ................................................................ 43

3.3 Previous impact models reviewed .......................................... 43

3.3.1 Example: Linear permanent impact (Bertsimas and Lo) ............ 44

3.3.2 Example: Temporary and permanent impact (Almgren and Chriss) 44

3.3.3 Example: Exponential propagator (Obizhaeva and Wang) .......... 45

4 Optimal execution with the propagator model ............................ 47

4.1 Statement of the problem ................................................... 47

4.2 Optimal execution .............................................................. 48

4.3 Optimal execution with spread costs ..................................... 49

4.4 Optimal execution with risk aversion ..................................... 50

4.4.1 Solution neglecting spread costs ...................................... 51

5 Estimation methods of the propagator model for market impact .... 53

5.1 Estimation of market impact of an individual transaction .......... 53

5.1.1 Adaptive binning algorithm ........................................... 54

5.2 Estimation of the propagator .............................................. 55

5.2.1 Estimation by linear regression ....................................... 55

5.2.2 OLS general solution ................................................... 55

5.2.3 Errors on the estimate .................................................. 56

5.2.4 Goodness of the regression ............................................ 57

6 Data and methods .................................................................. 59

6.1 Datasets description ........................................................... 59

6.1.1 2000 - 2002 dataset ...................................................... 59

6.1.2 2011 January - April dataset ............................................ 60

6.2 Software and libraries used .................................................. 63

7 Empirical estimate of propagator model ................................... 65

7.1 Different times for financial markets ...................................... 65

7.1.1 Trade time ....................................................................... 65

7.1.2 Aggregated trade time .................................................... 66

7.1.3 Real time ......................................................................... 66

7.2 Model in trade time ............................................................ 67

7.2.1 Volume dependence of market impact ................................ 67

7.2.2 Propagator ................................................................. 68

7.3 Model in aggregated trade time ............................................ 71
Introduction

In this Thesis we deal with two issues of great interest in the field of econophysics, the branch of physics that studies economy and finance: optimal execution and market impact. We briefly define them below.

We mention that this Thesis project started as a research collaboration with the Linear Quantitative Research Group of J.P. Morgan & Co. The group provided financial data, which we used to run our many empirical estimations, and invaluable insights into nature of the problems tackled.

We study the optimal execution of financial transactions, in particular transactions of equities. Many investors, like pension fund managers, need to trade substantial amounts of equities, far more than the market can offer at a given time. Thus, execution strategies were developed in order to split the original trade into many smaller trades, to be distributed over a period of time. In the last two decades new technologies have revolutionized the trading industry. The execution strategies have become increasingly refined: algorithmic trading, the usage of computer programs to manage trading, has emerged as an important practice in the finance industry.

Market impact is an effect that links transactions and price movements. It is observed that every trade on the market impacts the price: if the trade was initiated by a buyer the price moves up, if instead it was a seller the price moves down.

We will see that the modeling of market impact is crucial to determine a strategy for optimal execution. In general, a trading strategy requires the investor to split his original order in many small trades to be executed over a certain period of time. If each of those orders moves the price, all the subsequent trades will be affected, causing what are known as impact costs. By varying the temporal distribution of trades these impact costs can vary substantially. We will study how to build optimal strategies that minimize these costs.

A well known application of physical models in financial research is in the studying of stock price dynamics with the theory of brownian motion. A brownian particle hit by water molecules is compared to a stock hit by news that move its price.

The phenomenon of market impact enriches this picture, and has in fact been studied by many physicists. Trades as well as news move the price of a stock, and the way in which one models the relative contribution of the two effects is crucial. The interpretation proposed in \[1\], which is equivalent to the model we will work with, draws a very interesting picture. Think of a stock as a brownian particle in two dimensions, trades and news are molecules of two different kinds hitting it. One observes that the sign of the “momentum” of a trade molecule hitting the stock is a quantity strongly autocorrelated in time, i.e. a trade with momentum of a given sign is most probably followed by another whose momentum
has the same sign. News, on the other side, move in a truly random way. It then follows from a standard assumption in econophysics (unpredictability of price movements) that the effect of a trade hitting the stock depends on the history of trades. The brownian particle has some kind of memory of the past collisions, and reacts more strongly when hit by trades whose momentum has unexpected sign. In Chapter 1 we will explain the phenomenon of market impact in much greater detail.

The Thesis is composed of eight Chapters. The first three are mainly reviews of material found in the literature. The last five, instead, contain mostly original material. We briefly summarize their content.

• **1 - Efficient market hypothesis, random walk, and market impact** is an introductory Chapter. It reviews some important concepts of modern quantitative finance and econophysics. It starts by defining the stock market and the relevant variables we use to describe its dynamics. It then focuses on the modeling of the price series, first as random walks and then considering the contribution of market impact. In brief, market impact is the effect of trades on the price of a stock. If someone buys the stock the price rises. If on the contrary someone sells, the price drops.

• **2 - Review on optimal execution** covers the relevant literature (three seminal articles) on the problem. It is not merely a review, we reformulate the models in a common framework so that it is easier to compare them.

  Broadly speaking, optimal execution consists in finding a strategy to buy or sell a quantity of shares on the stock market, paying the least possible transaction costs. The cost of trading can be substantial, and market impact is an important component. We review three seminal articles which propose different models of market impact and therefore obtain different optimal trading strategies.

• **3 - Propagator model for market impact** introduces the market impact propagator formalism, proposed recently within the econophysics community (by Bouchaud et al.). It is a very flexible model for market impact. In the rest of the Thesis we will use it, and some develop new extensions. It generalizes the models seen in Chapter 2. It allows to estimate empirically the impact parameters from data of past market transactions.

• **4 - Optimal execution with the propagator model** contains the first original material we propose. It develops a theory for optimal execution in the market impact propagator framework. It aims at reproducing the optimal execution models reviewed in Chapter 2 in a new, much richer framework.

• **5 - Estimation methods of the propagator model for market impact** contains other original material. It explains a procedure for the empirical estimation of the parameters of the market impact propagator. The estimation is based on linear regression. We will obtain the impact parameters, an estimate of the associated errors and a measure of how well the model explains the data.
• **6 - Data and methods** examines the financial datasets we used for the subsequent empirical estimations. We explain how we handled and processed them, and applied some filtering procedures. It also briefly reviews the software, programming languages and libraries we used to perform our analysis.

• **7 - Empirical estimate of propagator model** shows our results for the empirical estimate of the market impact propagator parameters. It uses the theory developed in Chapter 5 on the data described in Chapter 6. We will study market impact in three different settings, corresponding to three definitions of time series. Two of them have never been attempted before. We will then compare the performances of the different models.

• **8 - Empirical calibration of optimal executions with the propagator model** studies the optimal execution strategies calibrated with the parameters obtained empirically in Chapter 7. We will analyze their costs, and compare them with the strategies reviewed in Chapter 2.

At last, we propose the Conclusions and some additional plots of trading strategies (Appendix A).
Acronyms

**EMH** Efficient Market Hypothesis.

**GARCH** Generalized Autoregressive Conditional Heteroskedasticity.

**IID** Independent and Identically Distributed.

**LOB** Limit Order Book.

**OLS** Ordinary Least Square.

**SSR** Sum of Square Residuals.
Chapter 1

Efficient market hypothesis, random walk, and market impact

This Chapter introduces notions and definitions which are well known in the econophysics community and literature. Most of the material covered here can be found in more details in a number of textbooks, as cited in the text.

We start by giving some basic concepts of modern quantitative finance, in sections 1.1 to 1.4.

Sections 1.5 and 1.6 instead focus on the problem of price formation and define market impact explaining how it has been recently treated in the econophysics literature. Most of the subsequent work of this Thesis builds on these ideas.

1.1 Financial markets

Financial markets allow two complementary populations to meet: entrepreneurs that have some industrial projects but need funding and investors that have money to invest. Investors will share, on the long run, both the future profits and the risk of these industrial projects.

There are many financial instruments traded on the markets, and many new ones are introduced every year. In this work we will only consider stocks, but many of the models we will develop could be extended to transactions of other financial assets.

1.1.1 Stocks

The stock of a company is, by definition, the original capital invested in the company by its founders. This stock is broken down into shares or equitites, which represent a fractional ownership (or equity) of the company.

The so called publicly quoted companies have stocks whose shares can be bought and sold freely in the financial markets. To buy shares of a privately held companies, instead, one must contact the owners directly and the transaction might be governed (when possible) by private ownership rules (see [2]).

We will focus on the dynamics of the transactions of stocks in the stock markets, therefore we will only consider stocks of public companies.
1.1.2 Stock markets

A stock market, or exchange, is an institution where stocks can be traded. It is characterized by an ecosystem of market participants and a set of regulations by which they interact. One can classify the market participants in three broad classes:

- **Investors** are individuals or companies (like mutual funds, see [2]) that buy shares of companies they wish to invest money in. They usually hold the stocks for months, or years: it is said that their investment has a long time horizons;

- **Intermediaries** are finance professionals that provide services for the other market participants. *Brokers* buy or sell stocks on behalf of a client, charging a fee. *Dealers* or *market makers* provide liquidity to the market by offering to buy or sell stocks at any given instant of time. They compensate for any momentary lack of counterparts that would make trading less liquid and riskier. They gain from the bid-ask spread, the difference between the price they ask to sell stocks (ask) and the price they offer to buy stocks (bid), as explained in section 1.1.3;

- **Speculators** take short to medium term positions on a stock (seconds to weeks). They essentially make bets on how the stock price will evolve.

This taxonomy is very common in the financial literature, but it is simplistic and in a sense outdated. In the last two decades, with the emergence of electronic markets, the difference between dealers and speculators has become more blurred. Also, the system by which most modern stock markets operate allows every individual to act as a market maker, by posting limit orders. The system is called continuous double auction, and it is explained in the next section.

1.1.3 A mechanism for markets: the limit order book

Almost all modern stock markets are operated by a Limit Order Book (LOB), a collection of all traders’ offers to buy and sell at different prices. Potential buyers notify the other participants of the maximum price at which they are willing to buy (bid price) a quantity of shares. Symmetrically, potential sellers declare the minimum price they require to sell (ask price) another quantity of shares.

The collection of all orders to buy and sell is called the order book, which is the list of total volume available for a trade at a given price. We show in figure 1.1 an example of the state of the LOB at a given instant of time.

In practice, every trader has two options:

- post a limit order, a commitment to buy/sell a certain volume of a stock at a price he decides. These are the orders that populate the LOB;

- instantly buy or sell using a market order. The investor specifies the volume of stocks he wants and the system matches his request with one or more limit orders in the book, starting from the best available price.
Figure 1.1: Example of a LOB for an exchange traded fund (analogous to a stock). The 15 best limit order are shown, both for buying and selling. For each order are displayed the number of shares offered and the buy or sell price. The two top quotes of each column are the best ask and best bid. Taken from [3].
For example, if the book is in the state of figure 1.1, a market order to sell 600 shares would match only the best buy limit order (600 shares at 25.124 $). If instead the market order was to sell 1000 shares, it would match the first buy limit order and part of the second. This means that the first 600 shares are sold at 25.124 $, and the subsequent 400 at 25.123 $.

We define the best ask price $A$ as the lowest price among the sell limit orders in the book. It is the lowest price that someone requires to sell some shares. Symmetrically, the best bid price $B$ is the highest price among the buy limit orders in the book. It is the highest price that someone offers to pay for buying some shares.

They are shown in figure 1.1. The left column contains the buy limit orders, the best bid is the price at its top. The right column contains instead the sell limit orders, the best ask is the price at its top.

We define the midprice $P$, which is the average of ask and bid: $P = \frac{A + B}{2}$. Also, the difference $s = A - B$ is called bid-ask spread.

### 1.2 Efficient Market Hypothesis

Market efficiency is one of the central ideas in finance. As a classic textbook on financial econometrics [4] put it, we can trace the origins of the Efficient Market Hypothesis (EMH) back to the pioneering theoretical contribution of Bachelier [5] and more recently Samuelson [6] and Fama [7]. Malkiel [8] has offered the following definition:

A capital market is said to be efficient if it fully and correctly reflects all relevant information in determining security prices. Formally, the market is said to be efficient with respect to some information set (...) if security prices would be unaffected by revealing that information to all participants. Moreover, efficiency with respect to an information set (...) implies that it is impossible to make profit by trading on the basis of [that information set].

We need to specify the information set with respect to which the market is efficient, i.e. which information set is incorporated in the security price. The classic taxonomy (due to Roberts [9]) is:

- **weak-form efficiency**: the information set includes only the history of prices or returns themselves;

- **semistrong-form efficiency**: the information set includes all information known to all market participants (*publicly available* information, such as the the history of prices, financial news, reports, etc.);

- **strong-form efficiency**: the information set includes all information known to any market participant (public and *private* information, such as research made by banks’ analysts for internal use).

There is strong consensus in the financial community that the strong form of EMH is not true. In fact, there are traders who have superior private information, and have clues about how the price will move.
1.3. PRICES, RETURNS AND RANDOM WALK

It is currently debated whether the weak and semi-strong forms hold, but they are nonetheless useful simplifications. The EMH asserts that financial markets are “informationally efficient”, so that the current price $P_n$ is the best predictor of future prices $P_{n+k}$:

$$E[P_{n+1}|P_n, \ldots, P_0] = P_n.$$  

(1.1)

This assumption, coupled with the requirement that the price series does not diverge, defines a martingale. Later (section 1.3) we will strengthen the requirements in order to model the price series with a special case of martingale: the random walk.

1.3 Prices, returns and random walk

We now turn to the dynamics of stock prices in the market. We consider the series $P_t$ of mid-prices of a stock, where the index $t$ runs over a specified time interval, say one second.

Virtually every aspect of quantitative finance involves returns rather than prices. The simple return on the stock between time $t$ and $t+1$ is defined as

$$R_t = \frac{P_{t+1}}{P_t} - 1.$$  

(1.2)

From this definition it is apparent that the stock’s return over the most recent $k$ periods (from time $t-k$ to time $t$), written $R_t(k)$, is given by:

$$R_t(k) \equiv (1 + R_t) \cdot (1 + R_{t-1}) \cdots (1 + R_{t-k+1}) - 1 = \prod_{j=0}^{k-1} (1 + R_{t-j}) - 1,$$  

(1.3)

which is called a compound return. The average return over the most recent $k$ periods is given by the geometric mean:

$$R_t^{avg}(k) \equiv \left[ \prod_{j=0}^{k-1} (1 + R_{t-j}) \right]^{1/k} - 1,$$  

(1.4)

a rather cumbersome mathematical relation. We now define another useful quantity, the logarithm of prices (known as log-prices):

$$p_t = \log(P_t).$$  

(1.5)

This definition will prove very useful in the following, and is widely used in the financial literature. One can also define log-returns:

\footnote{Here we are using a discrete time setting, whilst the price is in fact a continuous time process. We also note that there are many different definitions of discrete time in the stock market. In any case, the result we are going to show here are common to most of these definitions (at least the ones we will use in Chapter 7).}
\[ r_t \equiv \log(1 + R_t) = \log \left( \frac{P_{t+1}}{P_t} \right) = p_{t+1} - p_t. \]  

(1.6)

In most applications the returns \( R_t \) are small numbers, so that:

\[ r_t = \log(1 + R_t) \approx R_t. \]  

(1.7)

One of the advantages of using log-returns becomes clear when we consider multiperiod returns. With a little algebra equation 1.3 becomes:

\[ r_t(k) \equiv r_t + r_{t-1} + \cdots + r_{t-k+1}. \]  

(1.8)

Equation 1.4 for the average return translates instead in:

\[ r^{\text{avg}}_t(k) \equiv \frac{r_t + r_{t-1} + \cdots + r_{t-k+1}}{k}; \]  

(1.9)

which expresses the average log-return \( r^{\text{avg}}_t(k) \).

One of the earliest and most enduring questions in finance is whether financial assets returns, and stock returns in particular, are forecastable. The EMH, however, states that this is not possible. In mathematical terms this means that the stock price process is a martingale. We assume a stronger hypothesis, that the price series follows a random walk of independent and identically distributed increments, which we now define.

### 1.3.1 Arithmetic random walk

The strongest form of random walk assumes that the price increments are independent and identically distributed (IID) random variables. It is the first model we study, and the simplest. Assuming for simplicity no deterministic drift, the dynamics of price \( P_t \) is given by:

\[ P_t = P_{t-1} + \eta_t, \quad \eta_t \sim \text{IID}(0, \sigma^2) \]  

(1.10)

where IID(0, \( \sigma^2 \)) denotes that \( \eta_t \) are independently and identically distributed with average 0 and variance \( \sigma^2 \). This defines an arithmetic random walk.

We can express the conditional expectation \( \mathbb{E} \) and variance \( \text{V} \) of the price at time \( t \) given some initial value \( P_0 \) at time 0:

\[ \mathbb{E}[P_t|P_0] = P_0, \]  

(1.11)

\[ \text{V}[P_t|P_0] = \sigma^2 t. \]  

(1.12)

The most common assumption about the distribution of the increments \( \eta_t \) is normality. If the increments and distributed as \( \mathcal{N}(0, \sigma^2) \), then equation 1.10 is equivalent to an arithmetic Brownian motion, sampled at regularly spaced unit intervals.

There is a problem with the price dynamics model of equation 1.10. Under general assumption for the distribution of \( \eta_t \) (for example when they are normally distributed) the price \( P_t \) can become negative. Of course this cannot occur in reality and thus a different model has been proposed.
1.3.2 Geometric random walk

To solve the problems of arithmetic random walk, like the emergence of negative prices, one assumes that log-prices follow an arithmetic random walk:

\[ p_t = p_{t-1} + \eta_t, \quad \eta_t \sim \text{IID}(0, \sigma^2), \quad (1.13) \]

where the increments \( \eta_t \) are independent and identically distributed random increments. Recalling the definition 1.6 of log-returns we have that:

\[ r_t = p_{t+1} - p_t = \eta_{t+1}, \quad (1.14) \]

so the log-returns are distributed as IID(0, \( \sigma^2 \)). The variance \( \sigma^2 \) of the returns distribution is usually called volatility.

In case the distribution of \( \eta_t \) is normal, this model is equivalent to geometric Brownian motion sampled at regularly spaced unit intervals.

We note that the two random walk models between time \( t \) and time \( t + k \) are equivalent in the limit of \( P_t \gg \sigma \sqrt{k} \). When studying the price process over short periods of time we can thus use the two models interchangeably.

1.4 Stylized facts of financial time series

For most applications in finance researchers and practitioners use the geometric random walk model of equation 1.13, assuming normality of random increments \( \eta_t \).

However, this model fails to describe some statistical regularities of the empirical behavior of stock prices. We review some of the most interesting of these so called stylized facts.

1.4.1 Returns correlation

In the random walk hypothesis the returns are IID: it follows that they are linearly uncorrelated.

One defines the lagged autocorrelation function of the return series:

\[ \mathcal{C}(l) \equiv \frac{\langle r_k r_{k+l} \rangle}{\langle r_k^2 \rangle}, \quad (1.15) \]

where the notation \( \langle \ldots \rangle \) refers to the average over the whole return series.

The assumption that the stock log-returns \( r_t \) are IID implies that the autocorrelation function \( \mathcal{C}(l) \) is null for every lag \( l \neq 0 \). Figure 1.2 shows an example of an empirical autocorrelation function. The returns come from the Standard & Poor 500, a stock index. It is the average price, weighted by market capitalization (the total value of each company) of the 500 most capitalized American stocks.

We can see that for most of the points the autocorrelation is statistically consistent with zero, and therefore the returns are linearly uncorrelated for sufficiently long lags.\(^2\)

\(^2\)There are many other reasons, see [4].

\(^3\)This does not imply, however, that they are independent. For example, it has been proved that the squared returns are correlated on much longer time scales than the returns. See for example [10].
On a short time scale ($< 30$ minutes), however, significant correlations do exist. It is perhaps due to small statistical inefficiencies, because of market microstructure effects. It is interesting to note that this time scale keep shortening as the market technology progresses. In fact, it was of the order of hours in 1960, and seconds or less in 2011.

1.4.2 Fat tails

The assumption that returns are normally distributed is of great theoretical beauty and simplicity, but does not accurately describe the real stock price time series. As pointed out in many studies, see for example [11], [12] and [13], the discrepancy between model and reality is most notably seen in the tails of the distribution.

Figure 1.3 shows the cumulative distribution of empirical returns on the Standard & Poor 500. The empirical distribution (both for positive and negative return) is compared with some theoretical distributions.

Returns of high absolute value occur much more frequently than the Gaussian distribution would predict. In fact, the empirical return distribution is said to have fat tails. In mathematical terms such a distribution is called leptokurtic, because it has a kurtosis$^4$ higher than the Gaussian.

The other two theoretical distributions proposed in the figure model the empirical distribution much more precisely. Also, we note that the positive and negative returns distribution are not exactly symmetric. It is well known that extreme negative returns (crashes) are more likely to occur than extreme positive ones.

$^4$ The kurtosis of a zero-mean distribution of a random variable $x$ is defined as the fourth moment divided by the square of the variance, minus 3:

$$\gamma_1 = \frac{\mu_4}{\sigma^4} - 3 = \frac{E[x^4]}{E[x^2]^2} - 3.$$  \hspace{1cm} (1.16)

It equals 0 for a Gaussian distribution.
1.5. WHY DOES THE PRICE MOVE?

Figure 1.3: Empirical cumulative distribution of positive and negative returns of a stock index (Standard & Poor’s 500 over the period 1991-2001), compared to the Gaussian and two other theoretical distributions. The left graph shows the distribution of 30-minutes returns in log-log scale. The right graph shows instead the distribution of daily returns in a semi-log plot. Taken from [3].

1.4.3 Volatility clustering

The geometric random walk model assumes that the returns are IID random variables. In particular, this implies that their variance (or volatility), is constant throughout the series.

This is not true, as many studies have pointed out (see [14], [15]). The returns random process is heteroscedastic, i.e. its volatility varies with time - it is in fact a random process itself.

One can see in figure 1.4 the plot of the daily returns of a stock index (Dow Jones) over one century, compared to the returns of a random walk with the same variance.

It is evident from the figure that the return distribution shows a structure that is absent in the random walk returns. In particular, there are periods (clusters) when the volatility is high, and others when it is low. This phenomenon is called volatility clustering.

We should mention that in the field of financial econometrics many heteroscedastic models have been proposed to faithfully represent the real series. These are called stochastic volatility models, and perhaps the most used one is the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) [16].

1.5 Why does the price move?

In the previous section we have proposed and commented the random walk model for the price dynamics in the stock market. We have, however, left one fundamental question unanswered. Why does the price move, what causes the supposedly random movements
of the price process?
Two effects have been proposed to explain why the price moves: news and market impact.

1.5.1 Effect of news

The classic view on price dynamics asserts that the price moves because of the arrival of news on the market. If all market participants have access to the relevant information regarding the value of the stock (financial statements of the company, economic news, past prices, etc.) they should agree on a fundamental value\footnote{The reality is rather different. One major argument against this over-simplified model comes from studies that challenge the very idea of a “fundamental price”. Black \cite{black}, for example, argues that the intrinsic error on the evaluation of the fundamental value for a stock is as large as 100% the price.} for the price of the stock and trade accordingly. In this picture, the market price could change only when some unexpected news about the company come out. The market would then elaborate the news and agree on a new fundamental price, and keep it fixed until some other news come out. These news represent therefore the main “random” driver of the price process. This model, with the addition of some noise, is shown in figure 1.5.
1.5. WHY DOES THE PRICE MOVE?

Figure 1.5: The “classic view” of stock price evolution. The big movement of the price are due to the arrival of unexpected news that change the market consensus about the fundamental price of the stock. The price therefore adjusts for a mispricing $\Delta$.

1.5.2 Market impact

There is another driver of the price dynamics, a purely mechanical one which will prove to be very interesting in the rest of this Thesis. It has been proved empirically (see in particular [18], [1], [19], [20]) that transactions on the market move, on average, the midprices.

This effect is called market impact. In its original form it describes the market impact of individual transactions. In fact, it is observed that a transaction initiated by a buyer (with a buy market order) moves on average the midprice up. On the other side, a transaction initiated by a seller moves the midprice down.

We define the series $p^\mu_n$ as the logarithm of the midprice of the stock just before the $n$-th transaction on the market, and the returns $r^\mu_n = p^\mu_{n+1} - p^\mu_n$. We also define the series of volumes traded $v^\mu_n$ where $v^\mu_k$ is the volume of the $k$-th transaction. It is positive if the trade was initiated by the buyer (via a buy market order), and negative if it was initiated by the seller.

A question of particular interests, both for researchers and practitioners, is how the market impact of a transaction depends on its volume size. In [20] the problem has been studied empirically. It is found (figure 1.6) that the average log-return $r^\mu$ has a slowly increasing dependence on the volume $v^\mu$ of the corresponding transaction.

We now define the impact function $^7f(v^\mu)$. It is the expected return $r^\mu_n$ conditioned to a certain volume traded $v^\mu$ with its sign:

$$ f(v^\mu) = E[r^\mu|v^\mu]. \tag{1.17} $$

---

6 The superscript “tt” stands for trade time. This concept will be developed in greater detail in Chapter 7 section 7.1.1.

7 It will be defined more rigorously in Chapter 3 section 3.2.1.
Figure 1.6: Market impact of individual transactions as a function of the transaction volume, for different groups of stocks in the New York Stock Exchange, ordered by market capitalization (total value of the shares). The group labeled by “A” is the most capitalized. The absolute value of the log-price shift after a transaction $|\Delta p| = |v^t|$ is plotted against the volume $|v^t| \equiv \omega$ of the transaction. The results are nearly linear curves, with a positive slope clearly smaller then one. Taken from [20].

The curves in figure 1.6 are nearly linear in log-log scale. This means that the price impact is proportional to the volume elevated to some power $\lambda$, so that $f(|v^t|) \propto |v^t|^\lambda$. In the articles cited it is also found that the function $f(v^t)$ is odd, so that:

$$f(v^t) = \text{sign}(v^t) \cdot f(|v^t|).$$

(1.18)

We have therefore:

$$f(v^t) = A \cdot \text{sign}(v^t) \cdot |v^t|^\lambda$$

(1.19)

Various values of the coefficient $\lambda$ have been found in the literature, it is in any case a real number $\lambda \in [0, 1]$. The result of [20] is $\lambda = 0.3$. Some studies have proposed a logarithmic dependence $f(v^t) \propto \text{sign}(v^t) \cdot \log(|v^t|)$. In any case, the function $f(v^t)$ is odd.

With this result we can now express a basic model of geometric random walk with market impact:

$$p^t_n = p^t_{n-1} + \eta_n + f(v^t_n), \quad \eta_t \sim \text{IID}(0, \sigma^2)$$

(1.20)

where $\sigma^2$ is the variance of the stochastic component $\eta_n$ of return. In the next chapter we will use the impact of model of equation 1.20 with a simpler, linear $f$.

1.5.3 Order flow

We now turn our attention to an interesting empirical feature of the order flow hitting the book. Many works ([21], [18], [1]) have found that there is a significant temporal correlation of the signs $\epsilon_n = \text{sign}(v^t_n)$ of the market orders submitted in the book. We recall that a buy market order has positive sign, and a sell market order has negative sign.
Consider the series $\epsilon_n$ of the signs of the market orders submitted in the book. One can define the following autocorrelation function:

$$C(l) = \langle \epsilon_{n+l}\epsilon_n \rangle - \langle \epsilon_n \rangle^2. \quad (1.21)$$

If trades were random one should observe that $C(l)$ decays rapidly to zero, like the returns’ series autocorrelation. This does not happen: on the contrary, we can see in figure 1.7 that it decays very slowly to zero as an inverse power-law of $l$.

![Figure 1.7: Plot of the sign correlations](image)

The empirical autocorrelation is therefore modeled:

$$C(l) \simeq \frac{C_0}{l^\gamma}, \quad (l \geq 1). \quad (1.22)$$

We have that $C_0 \in [0, 1]$, because the signs are positively correlated. The value of $\gamma$ depends on the particular stock considered. It has been found in the literature in the range $0.2 \div 0.8$. It is smaller than one, therefore the order flow is said to be a long-memory process (see 23).

### 1.5.4 An apparent paradox

The picture we have just drawn leads us to an apparent paradox:

- we observed that each transaction, on average, moves the price of the stock in the direction of the trade. Equation 1.20 summarizes the idea;
we also saw that the signs $\epsilon_n$ of the order flow hitting the market are strongly autocorrelated in time. There is a positive probability that a market order will be followed by many others of the same sign.

We can show that from these two assumptions it follows that the returns series $r_n$ is positively autocorrelated, a fact that would strongly violate EMH and is not observed on the market. We start by expressing the return $r^t_n$, from eq. 1.20:

$$r^t_n = \eta_n + \epsilon_n f(|v^t_n|).$$  \tag{1.23}

Recalling that $\eta_n$ are IID random variables and $\langle \epsilon_n \rangle = 0$ we may write:

$$\langle r^t_{n+l} r^t_n \rangle = \langle \epsilon_{n+l} \epsilon_n \rangle \cdot \langle f(|v^t_{n+l}|) f(|v^t_n|) \rangle. \tag{1.24}$$

We now assume (as in [22]) that the last expression can be factorized:

$$\langle r^t_{n+l} r^t_n \rangle = \langle \epsilon_{n+l} \epsilon_n \rangle \cdot \langle f(|v^t_{n+l}|) f(|v^t_n|) \rangle. \tag{1.25}$$

We note that the the average of the impact functions is a positive number, and therefore we have positive autocorrelation of the returns. If in addition we assume that the last term is simply $\langle f(|v^t_n|) \rangle^2$ we have that:

$$\langle r^t_{n+l} r^t_n \rangle = \langle \epsilon_{n+l} \epsilon_n \rangle \cdot \langle f(|v^t_{n+l}|) f(|v^t_n|) \rangle = C_0 \cdot \langle f(|v^t_n|) \rangle^2 \cdot \frac{1}{l}, \tag{1.26}$$

as found in [22]. This has been proven empirically wrong \(^8\) (as we saw in figure 1.2), and also violates the EMH. Therefore, eq. 1.20 must be amended, it is not possible to have an impact which is fixed and permanent.

### 1.6 Consequences of correlated order flow

In order to solve the apparent paradox of section 1.5.4 a different model of price dynamics is proposed in [18], the impact propagator model. One assumes that the price can be written in general as:

$$p^H_n = p^H_0 + \sum_{k=0}^{n-1} \left[ G(n, k, v^H_k) + \eta_k \right] \tag{1.27}$$

where $\eta_k$ is a series of IID random shocks with mean 0 and variance $\sigma^2$, as usual. The function $G$ describes the impact at time $n$ of a trade at time $k$ of signed volume $v^H_k$. In [18] a key assumption was made (motivated by [24], [25]). The impact function $G$ is translation invariant, and its volume and time dependence can be factorised:

$$G(n, k, v^H_k) \equiv f(v^H_k) \cdot G_0(n - k). \tag{1.28}$$

The temporal part $G_0$ of the impact function is called the impact propagator.

---

\(^8\) The autocorrelation of returns that is observed for small lags is negative, while the order flow is positively correlated.
Bouchaud et al. [18] showed that by using an appropriate functional form for $G_0$ we can solve the “paradox” of section 1.5.4. If $G_0(l)$ decays with the lag $l$ it may in fact compensate for the long range correlations of the order flow, and restore the unpredictability of returns $r^\alpha$.

In [18] the exact mathematical relationship is stated. If the order flow autocorrelation decays with the lag $l$ as

$$C(l) \propto \frac{1}{l^\gamma},$$  \hspace{1cm} (1.29)

the price return is linearly uncorrelated if the propagator $G_0(l)$ decays as

$$G_0(l) \propto \frac{1}{l^{\beta}},$$  \hspace{1cm} (1.30)

with $\beta = \frac{1-\gamma}{2}$. We then obtain that the there is no predictability of returns $r^\alpha$, and the EMH is satisfied. In Chapters 3 and 7 we will develop and test various extensions of this model.

In this model the impact of a single transaction is fixed in value but decays with time. Another competing model was proposed in the econophysics literature (see [1]). As we briefly described in the Introduction, it assumes that the market impact of a trade depends on the trades’ history, i.e. a trade with unexpected sign impacts the price more than one whose sign is expected. In fact, as the authors on both parts have recently realized, the two models are equivalent [19].

The main objective of this Thesis is to use equation 1.27 to describe the price formation mechanism and study the problem of optimal execution. In fact, a big order of shares may need to be split in many smaller chunks, traded separately. It is therefore important to understand how market impact changes the price, leading to impact costs dependent on the execution strategy chosen.

In the next chapter we will describe the optimal execution problem and review some important model from the literature.
Chapter 2
Review on optimal execution

This chapter deals with the problem of \textit{optimal execution} by describing some of the most relevant models found in the literature. An optimal execution strategy is an algorithm to perform a sequence of stock transactions minimizing a relevant \textit{objective function}. In most cases the main concern is to minimize transaction costs, of which market impact is an important component. We will also see the minimization of a combination of costs and risk (to model \textit{risk aversion}).

We start, in section 2.1, by explaining why it is important to study optimal execution. We then concentrate on three seminal articles \cite{26,27,28}. We reformulate their models in a common framework, whose definitions are given in section 2.2. Section 2.3 describes the first model that emerged in the literature \cite{26}. Here the impact is assumed linear and permanent. The next two sections are based on \cite{27}: section 2.4 explains an extension to the previous model to add a \textit{temporary} component of impact; section 2.5 introduces the concept of risk aversion and optimizes a combination of costs and risk. Section 2.6 is dedicated to the model of \cite{28}, which proposed a transient, decaying model of market impact.

2.1 Motivations for optimal execution

In the last decades there has been a rapid growth in stock investing, driven by the increasing popularity of \textit{mutual funds}. This has led to a rising concentration of assets among institutional \textit{portfolio}\footnote{A \textit{portfolio} is a general range of investment. In particular, a stock portfolio is a set of stocks in which an investor has positions.} managers. An average fund can have substantial position in a stock, up to a few percent of the total market value of the company. Managers must frequently rebalance their portfolios, either to include new stocks, sell stocks, or to improve the tracking of a given index or benchmark. This generates many sizeable orders that must be executed in a relatively short period of time. In order to understand the problem we give some figures:

- the average volume of a stock traded in a market day ranges in $10^{-3} \div 10^{-2}$ of its market capitalization;
• the volume available in the book at a given time (sum of the volumes of limit order posted in the book) ranges in $10^{-5} \div 10^{-4}$ of the market capitalization;

• the volume an investment fund want to buy/sell can be as high as $10^{-2}$ of a company.

It is clear therefore that a fund manager can not just post a big market order for all the shares she wants to trade. Probably the volume offered in the book is not enough.

Also, she usually has strong concerns of possible information leakage. If a fund wants to enter or leave a big position in a company it probably has access to some additional information about the stock which must not be disclosed. Therefore, those big trades must be “concealed” somehow.

The common solution to this problem consists in an order splitting, as documented in many studies about the fund operations (see for example [29], [30], [31]). The investors break up their large trades into smaller “packages” that they execute over the course of several hours, if not days.

The transaction costs associated with trading such orders - often called execution costs - can be large. Execution costs comprise several components: explicit costs such as commissions or bid-ask spread and implicit costs, most notably the cost related to market impact, or impact cost.

Market impact is an unfavorable effect on prices that the very act of trading creates. As we saw in the previous Chapter, by the very act of selling the seller of a stock will push the price down, with decreasing yields as the sales proceed.

Finding an optimal execution strategy by minimizing execution costs can be a very complicated problem. It is of great interest for the fund managers and, in general, for big stock investors. In this Chapter we are going to review some of the most influential works on this topic.

2.2 Statement of the problem

We now build a framework for the problem of optimal execution, introducing the definitions that we will use in the rest of this work. Each of the articles we review in this Chapter uses slightly different sets of definitions and notations. By unifying the formalism we will make it simpler to appreciate the similarities and differences between the different models. We now define two objects of great importance for this work: the trading schedule, i.e. the program of trading that the investor follows, and its execution costs.

2.2.1 Trading schedule

Consider an investor who wants to trade a block of $X$ shares of a stock. If $X$ is positive (negative) the investor wants to buy (sell).

We fix a time horizon $T$ for the completion of the trading schedule. For simplicity, we assume that $T$ is a trading day at the London Stock Exchange, from 08:00 to 16:30.

We divide $T$ into $N$ intervals of length $\tau = \frac{T}{N}$, and call $t_k = k\tau$ the discrete times for $k = 0, 1, ..., N$. If we take $N = 102$ we have that each interval lasts 5 minutes, and the series of times is $t_k = 08:00$, 08:05, ..., 16:25, 16:30.
We define a *trading schedule* or *trading list*: \( v_0, v_1, ..., v_{N-1} \). It is a \( N - 1 \) entries vector \( v \), each element \( v_k \) is the volume traded between time \( t_k \) and \( t_{k+1} \); its sign is positive (negative) if the trade was initiated by a buy (sell) market order. We clearly have that:

\[
X = \sum_{j=0}^{N-1} v_j.
\]  

(2.1)

A *trading strategy* is now easily defined as a rule for determining the trading schedule \( v_k \) with the information available at time \( t_{k-1} \).

We note that all models we review in this Chapter determine trading schedules with constant signs, all equal to the sign of \( X \). If \( X \) is positive (negative) the investor has to buy (sell) shares and therefore uses buy (sell) market orders.

### 2.2.2 Execution costs

We now turn to the expressions of the execution costs for a trade list \( v_t \). We recall that \( P_t \) and \( p_t \) are the series of mid-prices and their logarithms.

We note that the price at which a share is traded at time \( k \) is not \( P_k \): commissions fees and, most notably, bid-ask spread raise this price if buying or lower it if selling.

We therefore define the series \( \tilde{P}_t \) which represent the *effective prices* at which shares are actually traded at every step. The relationship between \( \tilde{P}_t \) and \( P_t \) will be specified later.

As a simplification, one can for the moment assume that \( \tilde{P}_t = P_t \).

We can now express the costs resulting from trading along a certain trajectory. We define the execution or transaction costs \( C(v) \) as the difference between the total money paid or received during the execution (equal to the sum of volumes \( v_k \) traded at each step times the effective prices \( \tilde{P}_k \)) and the initial market value of the position:

\[
C(v) = \sum_{k=1}^{N} v_k \tilde{P}_k - XP_0.
\]

(2.2)

As we defined in eq. 2.2, execution costs are equivalent to the *implementation shortfall* introduced in [32], and they are better known in the finance industry as slippage.

We note that this expression for the execution costs need not to be positive, i.e. one may have negative costs (gains) from the execution. For example, if in a buying schedule the stock price declines substantially over the execution time horizon, \( C(v) \) could be negative (the total money paid is less than the initial market value of the position).

As a matter of fact, in all models we will study the price is a stochastic process. One therefore is not interested in the specific realization of the execution cost function, but its expected value \( E[C(v)] \). As we will see in the following, the variance of costs \( V[C(v)] = E[(C(v) - E[C(v)])^2] \) is useful, too.

### 2.3 Optimal execution with permanent linear impact

We start our review of optimal execution strategies with the model of Bertsimas & Lo [26], published in 1998. They used very simple market impact specifications. The price is
supposed to follow an arithmetic random walk, with the addition of a term of linear market impact:

\[ P_k = P_{k-1} + \theta v_{k-1} + \eta_{k-1}. \]  

(2.3)

Here \( \eta_k \) are IID random variables with zero mean and variance \( \sigma^2 \), and \( \theta > 0 \) is the market impact parameter. Therefore they suppose that the volume the investor trades at step \( k - 1 \) moves the price \( P_k \).

In this model one neglects the effect of bid-ask spread and the fees on execution costs. The effective price payed at the execution \( v_k \) at step \( k \) is assumed to be the next mid-price, so that:

\[ \tilde{P}_k = P_{k+1}. \]  

(2.4)

The execution costs are:

\[ C(v) = \sum_{k=1}^{N} v_k \tilde{P}_k - XP_0 = \sum_{k=0}^{N-1} (\eta_k + \theta v_k) \sum_{j=k}^{N-1} v_j \]  

(2.5)

We see that the effects of random price movements (\( \eta_k \)) and market impact (\( \theta v_k \)) are separated. Remembering that the random terms \( \eta_k \) are IID with zero mean one obtains the expected value of costs:

\[ E[C(v)] = \theta \sum_{k=0}^{N-1} v_k \sum_{j=k}^{N-1} v_j \]  

(2.6)

With a little algebra it becomes:

\[ E[C(v)] = \frac{\theta}{2} \left[ X^2 + \sum_{k=0}^{N-1} v_k^2 \right]. \]  

(2.7)

The variance of execution costs is instead:

\[ V[C(v)] = E[ (C(v) - E[C(v)])^2 ] = E \left[ \left( \sum_{k=0}^{N-1} \eta_k \sum_{j=k}^{N-1} v_j \right)^2 \right], \]  

(2.8)

here one needs to use the assumption that \( \eta_k \) are independent, so that \( E[\eta_i \eta_j] = 0 \) for \( i \neq j \). Therefore, if \( \sigma^2 \) is the variance of \( \eta_k \):

\[ V[C(v)] = \sigma^2 \sum_{k=0}^{N-1} (\sum_{j=k}^{N-1} v_j)^2. \]  

(2.9)

We note that if \( \eta_k \) are Gaussian random variables the distribution of \( C(v) \) is exactly Gaussian.
2.3. OPTIMAL EXECUTION WITH PERMANENT LINEAR IMPACT

2.3.1 Optimal solution

It is now easy to derive the trading list that minimizes execution costs. The problem consists in finding the trading list $v_t$ that minimize equation 2.7:

$$\min_v \mathbb{E}[C(v)].$$

(2.10)

In the original article [26] the authors used a complicated dynamic programming procedure to get the solution. In fact it suffices to look at equation 2.7 to note that the solution is:

$$v_k^{OPT} \equiv \frac{X}{N}, \frac{X}{N}, ..., \frac{X}{N}.$$  

(2.11)

The signs of $v_t$ are all equal to the sign of $X$. This solution consists in simply trading at a constant rate over the all period. Even if it may seem trivial, it is still a common practice in the trading industry [33].

The expected cost and variance of this solution, from equations 2.7 and 2.9, are:

$$\mathbb{E}[C(v_k^{OPT})] = \frac{1}{2} \theta \left( X^2 + \frac{X^2}{N} \right)$$

(2.12)

$$\mathbb{V}[C(v_k^{OPT})] = \frac{1}{3} \sigma^2 X^2 T \left( 1 - \frac{1}{N} \right) \left( 1 - \frac{1}{2N} \right)$$

(2.13)

This trajectory minimizes total expected costs, but the variance may be very large. We now express the trajectory which minimizes the variance of execution costs, a result that will be useful in section 2.5.

Minimum variance solution

The trading trajectory that bears minimum uncertainty on execution costs is given by minimizing equation 2.9. One therefore changes the objective function of the minimization problem:

$$\min_v \mathbb{V}[C(x_k)]$$

(2.14)

The new optimal trajectory $v_t^{MV}$ is computed very easily, observing the structure of eq. 2.9. It consists in selling the whole packet at the first step:

$$v_0 = X, \quad v_1 = ... = v_{N-1} = 0.$$  

(2.15)

So that:

$$\mathbb{E}[C(v_t^{MV})] = \theta X^2$$

(2.16)

$$\mathbb{V}[C(v_t^{MV})] = 0$$

(2.17)

This trajectory has null variance, but the expected execution cost can be as high as double the solution of eq. 2.11. In the section 2.5 following the work of Almgren & Chriss [27], we will review the trajectories that lie between these two extremes.
We recall that this model has some shortcomings: the \textit{transient} costs, for example bid-ask spread and commission, are not being accounted for; also, the effective trading price $\tilde{P}_k$ was assumed to be the mid-price at next step, $\tilde{P}_k = P_{k+1}$.

### 2.4 Optimal execution with temporary impact and spread costs

Almgren and Chriss \cite{27} proposed an interesting extension to the model of section \ref{2.3}, by adding a \textit{temporary impact} term and the bid-ask spread costs. In their model the effective trading price $\tilde{P}_k$ is now given by:

$$\tilde{P}_k = P_k + \text{sign}(v_k)\Delta + \rho v_k.$$  \hfill (2.18)

The term $\text{sign}(v_k)\Delta$ is the contribution from the bid ask spread. If fact, $\Delta$ equals half the bid-ask spread: in case of buying ($v_k$ positive) one pays $\Delta$ more than the mid-price, in case of selling ($v_k$ negative) one receives $\Delta$ less.

The term $\rho v_k$ represents a linear temporary impact. The assumption is that the market impact has a \textit{temporary} and a \textit{permanent} component. The temporary impact is more significant than the permanent one: most of the effect a trade has on price is assumed to be only transient. In the terminology of \cite{27}, the temporary impact accounts for the resilience of the limit orders in the book and its relaxation to a steady state after a temporary movement.

The equation \ref{2.5} for the execution costs becomes therefore:

$$C(v) = \sum_{k=0}^{N-1} v_k \tilde{P}_k - XP_0 = \sum_{k=0}^{N-1} (\eta_k + \theta v_k) \sum_{j=k+1}^{N-1} v_j + \sum_{k=0}^{N-1} (\text{sign}(v_k)\Delta + \rho v_k) v_k \quad \hfill (2.19)$$

And the expected value of the costs:

$$E[C(v)] = \frac{\theta}{2}X^2 + \left(\rho - \frac{\theta}{2}\right)\sum_{k=0}^{N-1} v_k^2 + \Delta \sum_{k=0}^{N-1} |v_k|, \quad \hfill (2.20)$$

If, as it is normally the case, all $v_k$ have the sign of $X$, the last term becomes $\Delta X$. Also, we note that if $\rho = \theta$ we obtain the same expected impact costs of the previous model (equation \ref{2.7}). Therefore, this model adds another degree of freedom to better tune the price impact to the actual empirical market behaviour.

The variance remains the same:

$$V[C(v)] = \sigma^2 \sum_{k=0}^{N-1} (\sum_{j=k}^{N-1} v_j)^2. \quad \hfill (2.21)$$
2.4.1 Optimal solution

The solution that minimizes equation 2.20 is the same that we found in section 2.3.1, as can be easily verified:

\[ v_{k}^{OPT} \equiv \frac{X}{N}, \frac{X}{N}, \ldots, \frac{X}{N}. \] (2.22)

However, the expected value of the cost changes:

\[ \mathbb{E}[C(v^{OPT})] = \frac{1}{2} \theta X^2 + \left( \rho - \frac{\theta}{2} \right) \frac{X^2}{N} + \Delta X. \] (2.23)

Therefore, adding a temporary impact term does not change the results of section 2.3.1. In fact, the optimal solution obtained is the same, only the expression of the expected cost 2.23 changes.

A central notion in modern finance is that investors are risk averse: given two investment opportunities with the same expected return, they will prefer the less uncertain one. A standard measure of uncertainty is the variance of the probability distribution of returns (or, in this case, costs).

2.5 Risk aversion and efficient frontier of optimal execution

Following [27] we review another extension to the model of section 2.3. It builds on the concept of risk aversion, and defines a whole class of optimal solutions, one for every level of risk aversion.

Consider the whole class of possible trading schedules \( v_t \). Almgren and Chriss draw a cost versus variance plot: a cartesian coordinate system where the \( x \) axis represents the costs’ variance \( \mathbb{V}[C(v)] \) and the \( y \) axis the expected value of costs \( \mathbb{E}[C(v)] \). For each possible schedule one draws the corresponding point in the plot.

The result is a whole region of allowed values, whose border contains efficient frontier of optimal solutions. It is shown in Figure 2.1.

Almgren and Chriss define the efficient frontier as the set of solutions with minimum execution costs, given a maximum level of variance allowed\(^2\).

For each level of variance \( \Lambda \), one takes the schedule \( v_{t}^{*} \) that minimizes costs:

\[ \min_{v_{t} : \mathbb{V}[C(v)] \leq \Lambda} \mathbb{E}[C(v)], \quad \Lambda > 0. \] (2.24)

With this definition we have indexed a whole class of solutions with the parameter \( \Lambda \).

Definition with Lagrange multiplier

The definition of efficient schedules as solutions of the constrained optimization problem (2.24) may be improved. The schedules are parametrized by the variable \( \Lambda \), representing the maximum allowed level of variance of execution costs.

\(^2\) Or alternatively, the set of solutions with minimum variance given a maximum level of costs allowed.
Figure 2.1: The gray region is the set of all points representing the variance and expected cost of some possible trading schedule. The solid curve, part of the region’s border, is the *efficient frontier* of optimal execution. The solution of letter B is the minimum costs schedule, like that of eq. [2.11]. The solution of letter A would instead be chosen by a risk-averse investor: at the expense of higher costs the variance is lower. Letter C shows a solution that is not part of the efficient frontier, with high variance. An hypothetical *risk seeking* investor, one that is willing to pay more in order to take more risk, would choose it. Taken from [27].
One instead may introduce a Lagrange multiplier $\lambda$, and $2.24$ becomes an unconstrained minimization problem:

$$
\min_{v_t} \left( E[C(v)] + \lambda V[C(v)] \right).
$$

(2.25)

The Lagrange multiplier $\lambda$ indexes the whole boundaries of the region shown in Figure $2.1$. In particular, optimal solutions are the ones obtained with $\lambda \in [0, \infty]$, negative values of $\lambda$ correspond instead to the non-optimal dashed line of Figure $2.1$. By assigning extreme values to the parameter $\lambda$ one obtains the two results of section $2.3.1$:

- $\lambda = 0$ means that the investor is risk neutral, he does not care about risk. The solution therefore minimizes expected cost, such as in equation $2.11$.
- if instead $\lambda = +\infty$ the investor is extremely risk averse. He does not care how much more he has to pay, he just wants to minimize risk. One therefore obtains the solution with minimum (null) variance of equation $2.15$.

It is interesting now to study the solutions that lie between these two extreme values.

### 2.5.1 Optimal solutions for different parameters of risk aversion

In this section we report Almgren and Chriss’ general solutions to the minimization problem of equation $2.25$. They use the “temporary impact” model: the expression $2.20$ and $2.21$ for the expected execution costs and variance.

The solutions, parametrized by the Lagrange multiplier $\lambda$, can be obtained analytically (see $[27]$ for the details):

$$
v_k = A \cosh(\beta(T - t_k)),
$$

(2.26)

where $A$ is a normalization constant. The coefficient $\beta$ is related to the coefficient of risk-aversion $\lambda$, and is given by:

$$
\beta = \sqrt{\frac{\lambda \sigma^2}{\rho}}, \quad \lambda > 0.
$$

(2.27)

The equation found in $[27]$ is slightly different because of the way the temporary impact and the variance, proportional to the interval length $\tau = T/N$. The results are in any case equivalent.

These are the optimal solutions that form the efficient frontier. We will use the functional form of this solution in our numerical simulations, in particular in section $8.4$.

In figure $2.2$ we plot the trading trajectories $v_t$ for various value of the parameter $\beta$. We can see that by increasing the parameter $\beta$, and therefore the risk aversion $\lambda$, one gets solutions which are more and more front-loaded. Most of the trading is done at the start of the period, so that the investor is less exposed to the price fluctuations.
Figure 2.2: Almgren & Chriss’ optimal solutions, as in equation 2.26. We plot, for various values of the parameter $\beta$ (related to the risk-aversion $\lambda$) the corresponding trading lists $v_t$.

### 2.6 Optimal execution with exponentially decaying impact

We are now going to review the work of Obizhaeva & Wang [28]. The basic idea that motivates their work is that a trade impacts the market price by changing the stock’s current supply/demand. In addition, a change in current supply/demand can lead to evolutions in future supply/demand, which will affect the costs for future trades. In other words, in order to understand the market impact and its evolution with time, they model the supply and demand of stocks in the market. i.e. the effect that trades have on the book, depleting limit orders and widening the bid-ask spread. We will recast their impact model in our formalism, by concentrating on the movement of the mid-price following a trade.

Every trade has two effects on the price: a permanent impact term, like the one we saw in section 2.3, and a transient impact which is exponentially decaying with time. The impact $I$ on the price of a trade $v_k$ is therefore modeled as:

$$I(v_k) = \theta v_k + \rho v_k e^{-\gamma t},$$  \hfill (2.28)

with $\theta$ is the permanent impact parameter, $\rho$ the transient impact and $\gamma$ the decay exponent, all positive reals. The impact decays with time, therefore a trade taking place now affects the price changes for many steps in the future.

Adding the usual random shocks $\eta_k$, one has that the mid-price $P_k$ at time $k$ is given by:
2.6. OPTIMAL EXECUTION WITH EXPONENTIALLY DECAYING IMPACT

\[ P_k = P_0 + \sum_{j=0}^{k-1} \theta v_j + \sum_{j=0}^{k-1} \eta_j + \sum_{j=0}^{k-1} \rho v_j e^{-\gamma(t_k-t_j)}. \]  

(2.29)

In the model, the effective price of the trade at step \( k \) is expressed as:

\[ \tilde{P}_k = P_{k+1} + \Delta \text{sign}(v_k), \]  

(2.30)

where \( \Delta \) is again half the bid-ask spread.

Therefore, the expected cost of trading of a given schedule \( v_t \) is:

\[
E[C(v)] = \theta \sum_{k=0}^{N-1} v_k \sum_{j=k+1}^{N-1} v_j + \rho \sum_{k=0}^{N-1} \sum_{j=0}^{k-1} v_j v_k e^{-\gamma(t-k-j)} + \Delta \sum_{k=0}^{N-1} v_k. \]  

(2.31)

The term in \( \theta \) is linear permanent impact, as in eq. (2.6). The second term is generated by the transient exponential evolution of the price impact, while the third is due to the bid-ask spread costs.

2.6.1 Optimal solution

Minimizing the expected transaction costs function of eq. (2.31) one obtains the optimal solution. The calculation is rather lengthy and complicated, see [28] for the details. Figure 2.3 shows a plot of the solutions arising from the model, for different numbers \( N \) of steps in which the trading time horizon is divided.

![Trade Profiles for Different N](image)

Figure 2.3: Optimal solutions for the exponentially decaying impact model. The trade lists \( v_t \) that minimize the execution costs of eq. (2.31) are shown. The trade size \( X \) is fixed at \( 10^6 \) shares, and the resilience parameter is \( \gamma = 2.31 \ T^{-1} \). Taken from [28].
The functional form of the solution \( v_t \) is:

\[
v_0 = v_{N-1} = \frac{X + \gamma TX/N}{\gamma T + 2}, \quad v_k = \frac{\gamma TX/N}{\gamma T + 2}, \quad k = 1, 2, \ldots, N - 2.
\] (2.32)

Qualitatively, we can see that this trading schedules consists in two seemingly unconnected patterns:

- trading big chunks of shares, at the very start and end of execution;
- holding a constant trading rate at a much lower level for the rest of the time horizon, similarly to the “flat” solution of eq. 2.11.

These two different paradigms are due to the transient and decaying nature of impact: the parameter \( \gamma \) is crucial in determining their relative importance.

If \( \gamma T \gg 1 \) the transient impact decays very rapidly. In the limit of \( \gamma T \to \infty \) one recovers the model of section 2.4: the exponential decaying impact becomes a temporary impact, one can verify that the optimal solution converges to eq. [2.22].

Instead, if \( \gamma T \) is small compared to 1, the impact decays very slowly over the trading period. It is therefore convenient to separate as much as possible in time (at the very start and very end of the trading period) most of the shares traded, to let the exponential impact decay and to pay less impact costs.

This exponentially decaying impact model yields an optimal solution (minimal execution costs schedule) very different in nature than the solution 2.11 of the model with impact constant in time. Notice that both models are rather crude. It would be interesting to derive and assess optimal trading strategies based on more realistic impact models.
Chapter 3

Propagator model for market impact

This Chapter is dedicated to the *market impact propagator* formalism, which we outlined at the end of Chapter 1. It was proposed by econophysicists, in the 2004 paper by J.P. Bouchaud, Y. Gefen, M. Potters and M. Wyart [18].

Section 3.1 explains the reasons to study this new formalism, addressing the shortcomings of the impact models we reviewed in the last Chapter.

Section 3.2 defines the impact propagator model and the *impact function*.

In section 3.3 we show that the propagator formalism can accommodate the impact models reviewed in Chapter 2; this section is original.

### 3.1 Motivations for the market impact propagator model

We have seen in Chapter 2 some of the most significant approaches to the problem of finding an optimal trading strategy for the execution of large stock transactions. Market impact plays an important role in those models. We have seen that a market impact permanent in time (sec. 2.3) and another exponentially decaying (sec. 2.6) give rise to fundamentally different optimal solutions.

However, we argue that the literature we reviewed is not satisfying on the specific issue of market impact modeling:

- Bertsimas and Lo [26] assume an impact model constant in time, a very simplifying assumption that we have shown to be inconsistent with empirical findings in Chapter 1 (the “apparent paradox” of section 1.5.4). We noted that because of the strong autocorrelation of the signs of the orders hitting the market, one cannot have a market impact fixed and permanent: it would generate autocorrelated returns that strongly violate the efficient market hypothesis. Models with market impact that *decays* in time are needed;

- Almgren and Chriss [27] add a temporary component to the market impact which disappears right after the next transaction. Their model attempts to give a temporal structure to the market impact, but still it is too simplistic: the transition between “temporary” and “permanent” regimes must be smooth;
Obizhaeva and Wang [28] postulate a fully dynamic modeling of impact, exponentially decaying, somehow addressing the problems with the other models. However, they do not justify empirically the assumption that impact decay is exponential.

The problem of market impact has been studied thoroughly in the econophysics literature, from a very different perspective. In particular, [18], [1] and [19] bring in a new formalism, the market impact propagator, that we outlined at the end of Chapter 1.

The most interesting feature of this formalism is that we do not need to make any assumption on the particular functional form of the temporal structure of market impact. In fact, the model allows to completely estimate it empirically from publicly available data on past market transactions.

In the rest of this Chapter we review the propagator formalism, and re-express the models of Chapter 2 in this new framework.

### 3.2 Definition of the propagator model

We define a series of discrete times $t_n$, indexed by $n \in \mathbb{N}$. The price $p_k$ is the log mid-price of the stock at, or right before, time $t_k$. The series of volumes traded on the stock $v_k$ is the algebraic sum, for all trades between time $t_k$ and time $t_{k+1}$, of the volumes of the individual transactions (each signed as positive if buyer initiated and negative if seller initiated). The propagator formalism was originally developed by Bouchaud et al. in [18] with a particular choice of discrete time series, known as trade time $t^1$.

We anticipate that we are going to generalize their construction, allowing for different choices of time series $t_n$, from which follow the price and volume series $v_n$ and $r_n$. In Chapter 7 we will use these series to estimate empirically the market impact propagator.

One introduces the general impact function $G$, which describes the impact at time $t_j$ of the trade of volume $v_k$ between time $t_k$ and time $t_{k+1}$. It is nonzero only for $j > k$, because of causality.

In [18] a key assumption was made. The volume and time dependence of the general impact function $G$ can be factorised:

$$
G(j,k,v_k) \equiv f(v_k) \cdot G(j,k).
$$

In addition, one can assume (as found in [24] and [25]) that the market impact function is translation invariant, so that the market impact propagator $G_0$ can be defined:

$$
G(j,k) \equiv G_0(j - k).
$$

---

1. They assumed that each element of $t_n$ corresponds to the time of a transaction on the stock. It then follows that $p_k$ is the log midprice just before the $k$-th transaction, and $v_k$ the number of shares traded in the $k$-th transaction, signed as positive (negative) if the trade was initiated by the buyer (seller). This choice is a concept that we will develop in greater detail in Chapter 7, section 7.1.1.
3.3. PREVIOUS IMPACT MODELS REVIEWED

It models the temporal evolution of impact. Recalling that $G$ is non-zero only for $j > k$ we have that $G_0(t)$ is non-zero only for $t > 0$.

One therefore models the price dynamics by a geometric random walk with market impact:

$$p_j = p_0 + \sum_{k=0}^{j-1} [\eta_k + f(v_k)G_0(j - k)],$$

(3.4)

where $\eta$ is a series of IID random variables of mean 0 and variance $\sigma^2$. One may write explicitly the difference between two consecutive prices (log-return):

$$r_j = p_{j+1} - p_j = \eta_j + G_0(1)f(v_j) + \sum_{k>0} [G_0(k + 1) - G_0(k)]f(v_{j-k}).$$

(3.5)

The return is made up of three components: a random shock $\eta_j$, the effect of the decay of the impacts of all older transactions $\propto f(v_{j-k})$, and the impact of the last transaction $\propto f(v_j)$.

### 3.2.1 Impact function

By the factorization of the volume and time dependence of market impact one obtains the instantaneous “impact function” $f(v)$.

This function is defined as the expected return $r_n \equiv p_n - p_{n-1}$ conditioned to a certain volume traded $v$ with its sign.

$$f(v) = E[r|v],$$

(3.6)

This function has many interesting properties. It is assumed to be odd, so that $f(-v) = -f(v)$. We recall that the volume $v$ has positive sign if the trades were buyer initiated, and negative if seller initiated. We will see in the estimations of chapter 7 that this antisymmetry is well verified. In the following we may write: $f(v) = \text{sign}(v)f(|v|)$.

Also, many studies ([18], [1], [19]) have pointed out that $f$ is a concave function of the volume $v$, i.e. one that increases rapidly for small $v$ and more slowly for larger $v$.

It can be estimated on a series of financial data in a very simple fashion, following [22]. One computes the expected value of eq. 3.6 by averaging over all the sample. In Chapter 5 we explain in detail the methods we use to estimate this function, and in Chapter 7 we show the results of our estimations.

Now instead we show how the propagator formalism can accommodate the impact models reviewed in Chapter 2.

### 3.3 Previous impact models reviewed

The propagator formalism generalizes the impact models of Bertsimas and Lo, Almgren and Chriss and Obizhaeva and Wang, which we studied in the last Chapter. It is now easy to reformulate them.
CHAPTER 3. PROPELLER MODEL FOR MARKET IMPACT

3.3.1 Example: Linear permanent impact (Bertsimas and Lo)

We start by modeling the linear and permanent impact of Bertsimas and Lo, reviewed in section 2.3. The impact function is given by \( f(v) = \theta v \).

The price \( P_n \) is given by linear combination of past transactions:

\[
P_n = P_0 + \sum_{k=0}^{n-1} [\eta_k + \theta v_k G_0(n-k)].
\] (3.7)

We now set \( G_0(k) = 1 \ \forall \ k > 0 \), and recover Bertsimas and Lo’s impact model:

\[
P_n = P_0 + \sum_{k=0}^{n} [\eta_k + \theta v_k].
\] (3.8)

3.3.2 Example: Temporary and permanent impact (Almgren and Chriss)

Almgren and Chriss added to the model of Bertsimas and Lo a term of temporary market impact. It is defined as the impact that affects the price of the current transaction, but suddenly disappears afterwards.

However, we have seen that the impact propagator \( G_0 \) has a null term \( G_0(0) \) because of causality. We thus introduce the effective propagator \( \tilde{G}_0 \):

\[
\tilde{G}_0(t): \mathbb{N} \to \mathbb{R},
\] (3.9)

it is defined in such a way that the effective price, which is directly involved in the estimation of transaction costs, is written as:

\[
\tilde{P}_n = P_0 + \sum_{k=0}^{n} [\eta_k + f(v_k) \tilde{G}_0(n-k)].
\] (3.10)

In this case, it extends the propagator \( G_0 \) with the term \( \tilde{G}_0(0) \), which represents the instantaneous impact.

In this fashion, we have that the volume \( v_k \) affects the effective price \( \tilde{P}_k \) which one pays on the execution of the same \( v_k \), a feature that cannot be modeled with the propagator \( G_0 \) because \( G_0(0) = 0 \).

The instantaneous impact of a transaction is linear, with coefficient \( \rho > 0 \). The impact function is therefore given by \( f(v) = \rho v \). The permanent impact is also linear, with coefficient \( \theta > 0 \).

The effective price \( \tilde{P}_n \) is given by (neglecting for the moment the bid-ask spread contribution):

\[
\tilde{P}_n = P_0 + \sum_{k=0}^{n} [\eta_k + \rho v_k \tilde{G}_0(n-k)].
\] (3.11)

\[\text{\footnotesize\textsuperscript{2}}\text{ Here we are modeling an arithmetic random walk, while the propagator formalism has been developed with geometric random walk. We should not worry, because in normal conditions (trading time horizons shorter than a few days) the two random walk models are equivalent.} \]
3.3. PREVIOUS IMPACT MODELS REVIEWED

Setting $\tilde{G}_0(0) = 1$, and $\tilde{G}_0(k) = \frac{q}{\rho} \forall k > 0$ we obtain the model of Almgren and Chriss:

$$\tilde{P}_n = P_0 + \sum_{k=0}^{n-1} [\eta_k + \theta v_k] + \eta_n + \rho v_n. \quad (3.12)$$

3.3.3 Example: Exponential propagator (Obizhaeva and Wang)

The last model we reviewed in Chapter 2 is the richest, with a full temporal structure of market impact. The impact function is again linear, given by $f(v) = (\rho + \theta)v$.

The impact propagator has instead a more interesting structure. Calling $\tau$ is the time between $t_n$ and $t_{n+1}$:

$$\tilde{G}_0(k) = \frac{\theta}{\rho + \theta} + \frac{\rho}{\rho + \theta} e^{-\gamma k \tau}. \quad (3.13)$$

Again, $\tilde{G}_0(0) = 1$, and the effective price is given by:

$$\tilde{P}_n = P_0 + \sum_{k=0}^{n} \eta_k + \sum_{k=0}^{n} (\rho + \theta) v_k \tilde{G}_0(n - k) \theta v_k =$$

$$= P_0 + \sum_{k=0}^{n} \eta_k + \sum_{k=0}^{n} [\theta v_k + e^{-\gamma (n-k) \rho v_k}] \quad (3.14)$$

which is the same as eq. 2.29.

As anticipated, the propagator formalism allows to recast in a very compact form the models we previously introduced.
Chapter 4

Optimal execution with the propagator model

In this chapter we develop the models we will use to build optimal execution strategies that minimize market impact (with our $G_0$ model). It will be applied in Chapter 8. All material in this chapter is original.

We start in section 4.1 recalling some of the definition used in Chapter 2. Section 4.2 describes the basic model of optimal execution in the market impact propagator framework. Section 4.3 extends the model by adding the cost of bid-ask spread. Section 4.4 instead deals with the problem of risk aversion, and how it affects the optimal execution.

4.1 Statement of the problem

We recall some of the more relevant definitions we gave in Chapter 2. First, the trading schedule $v$ of section 2.2.1. Every element $v_k$ of the series is the volume we trade between time $t_k$ and time $t_{k+1}$, positive if we buy and negative if we sell. We have that $\sum_{k=0}^{N-1} v_k = X$. The effective price $\tilde{P}_k$ is the price at which we trade in the $k$-th step. The execution costs $C(v)$ were defined in section 2.2.2 as:

$$C(v) = \sum_{k=0}^{N-1} v_k \tilde{P}_k - X P_0 = \sum_{k=0}^{N-1} v_k (\tilde{P}_k - P_0).$$  (4.1)

We now define the logarithmic transaction costs $c(v)$, where $\tilde{p}_k$ is the log of the effective price $\tilde{P}_k$:

$$c(v) \equiv \sum_{k=0}^{N-1} v_k (\tilde{p}_k - p_0).$$  (4.2)

We note that in the limit of short time horizons of the execution $T$, the effective price $\tilde{P}_k$ does not deviate much from the starting value $P_0$. We can thus approximate:

$$c(v) = \sum_{k=0}^{N-1} v_k \log \left( \frac{\tilde{P}_k}{P_0} \right) \simeq \sum_{k=0}^{N-1} v_k \left( \frac{\tilde{P}_k - P_0}{P_0} \right) = \frac{C(v)}{P_0}. \quad (4.3)$$

47
The logarithmic execution costs $c(v)$ can therefore be thought as fractional execution costs. One important thing follows: to minimize the actual execution costs $C(v)$ it is equivalent to minimize the logarithmic execution costs $c(v)$.

We also define the fractional execution costs per share $\bar{c}(v)$:

$$\bar{c}(v) \equiv \frac{\sum_{k=0}^{N-1} v_k (\hat{p}_k - p_0)}{|\sum_{k=0}^{N-1} v_k|} = \frac{c(v)}{|\sum_{k=0}^{N-1} v_k|}. \quad (4.4)$$

We are now ready to express the execution costs within the framework of the market impact propagator developed in Chapter 3. We start from equation 3.4, which we report here:

$$p_n = p_0 + \sum_{k=0}^{n-1} [\eta_k + f(v_k)G_0(n-k)]. \quad (4.5)$$

To model the effective price $\hat{p}_n$ we also recall the definition of the effective propagator $\hat{G}_0(t)$ defined in section 3.3.2, which has a non-zero instantaneous impact term $\hat{G}_0(0)$. We have therefore:

$$\hat{p}_n = p_0 + \sum_{k=0}^{n} [\eta_k + f(v_k)\hat{G}_0(n-k)]. \quad (4.6)$$

## 4.2 Optimal execution

We start with a basic model of optimal execution within the impact propagator $G_0$ model. We are in a sense following the first step of Chapter 2, where we first reviewed the simple model of Bertsimas and Lo (section 2.3), and then analyzed its extensions.

From equation 4.6 follows that the fractional execution costs $c(v)$ are given by:

$$c(v) \equiv \sum_{k=0}^{N-1} v_k (\hat{p}_k - p_0) = \sum_{n=0}^{N-1} v_n \left[ \sum_{k=0}^{n} (\eta_k + f(v_k)\hat{G}_0(n-k)) \right]. \quad (4.7)$$

We are interested in the expected value:

$$E[c(v)] = \sum_{n=0}^{N-1} v_n \left[ \sum_{k=0}^{n} f(v_k)\hat{G}_0(n-k) \right] \quad (4.8)$$

We proceed, like in Chapter 2 by assuming that $f$ is linear\footnote{For example we will see in sections 7.3.1 and 7.4.1 that the impact function for aggregated trade time and real time is very close to be linear.}. We define a series $\theta_n$ of coefficients such that:

$$f(v_k) = \theta_k v_k. \quad (4.9)$$

In the following we assume that the elements are constant, $\theta_n \equiv \theta$. We define the $(N-1) \times (N-1)$ triangular impact matrix:

$$I_{i,j} = \begin{cases} \theta_i \hat{G}_0(i-j) & i \geq j \\ 0 & i < j \end{cases}. \quad (4.10)$$
4.3. **OPTIMAL EXECUTION WITH SPREAD COSTS**

We now have a matrix form for the execution costs expression:

\[ E[c(v)] = v^T I v. \]  

(4.11)

We want to find the optimal trading schedule, the one that minimizes:

\[ \min_v v^T I v, \]  

(4.12)

subject to the constraint, calling \( \mathbf{1} \) a vector whose \((N-1)\) elements are all 1:

\[ \sum_k v_k = \mathbf{1}^T v = X. \]  

(4.13)

This a constrained minimization problem, analogous to the well known Markowitz best portfolio problem (see [34], [35]).

We use a Lagrange multiplier \( z \) to enforce the constraint. We obtain therefore the following unconstrained minimization problem:

\[ \min_v \Lambda(v, z) = \min_v (v^T I v + z \cdot (\mathbf{1}^T v - X)). \]  

(4.14)

We set the derivative along \( v \) equal to zero:

\[ \frac{\partial \Lambda(v, z)}{\partial v} = 2Iv + z\mathbf{1} = 0 \implies I v = -\frac{z}{2} \mathbf{1}. \]  

(4.15)

The optimal \( v^* \) is:

\[ v^* = -\frac{z}{2} I^{-1} \mathbf{1} = \frac{X}{\mathbf{1}^T I^{-1} \mathbf{1}} I^{-1} \mathbf{1}. \]  

(4.16)

And its expected execution costs:

\[ E[c(v^*)] = v^{*T} I v^* = X^2 \frac{(\mathbf{1}^T I^{-1} \mathbf{1})(\mathbf{I}^{-1} \mathbf{1})}{(\mathbf{1}^T I^{-1} \mathbf{1})^2} = \frac{X^2}{\mathbf{1}^T I^{-1} \mathbf{1}}. \]  

(4.17)

We note that the execution costs are quadratic in the total volume \( X \), the same result we obtained in section 2.3. The most interesting feature of the model we just studied is that the optimal trading trajectory \( v^* \) can be computed with a simple closed-form matrix expression. This result is similar to the one obtained in [36].

This result is of great appeal. We saw while reviewing Obizhaeva & Wang exponential impact model in section 2.6 that increasing the impact model complexity can make it very difficult to find the optimal schedule. With our approach, even the most complex impact form \( G_0 \) can lead to very simple optimal solution \( v^* \). However, the picture gets a little more complicated if we include the contribution of bid-ask spread costs.

### 4.3 Optimal execution with spread costs

One important extension to the model of last section is to consider the contribution of bid ask spread costs. In doing this, we are loosely following Chapter 2. In fact, section 2.4 added to the basic model of Bertsimas and Lo the contribution of bid-ask spread and a better modeling of effective price \( \tilde{p}_k \).
We therefore follow ?? and assume that the bid-ask spread has a constant value $s = A - B$. Recalling that $P = \frac{A+B}{2}$ is the stock midprice we define $\delta$, half the fractional bid-ask spread:

$$\delta \equiv \frac{s}{2} = \frac{A - B}{A + B}. \tag{4.18}$$

Then, mimicking equation 2.18 we express the effective log-price $\tilde{p}_n$ as:

$$\tilde{p}_n = p_0 + \sum_{k=0}^{n} [\eta_k + f(v_k)\tilde{G}_0(n-k)] + \text{sign}(v_k)\delta, \tag{4.19}$$

because we pay half the bid-ask spread on execution. The expected value of the transaction costs including the bid-ask spread contribution is therefore:

$$E[c(v)] = v^T\mathcal{I}v + \delta 1^T|v|. \tag{4.20}$$

Minimizing 4.20 is more complicated then 4.11. This objective function is not linear, because of the term with the absolute value of $v$. We can not use anymore the machinery of Lagrange multipliers and matrix inverse (eq. 4.16).

As we will see in sections 8.2 and 8.3, when we will apply this theory, we will have to use numerical minimization methods.

### 4.4 Optimal execution with risk aversion

In this section we propose an extension of the optimal execution model, to accommodate a factor of risk-aversion. The idea that risk matters in determining the optimal execution schedule was first proposed by Almgren & Chriss in [27]. In section 2.5 we have reviewed their construction.

We adapt their approach to our model. We express the variance $\mathbb{V}[c(v_t)]$ of the fractional transaction costs $c(v_t)$ of equation 4.7.

We assume that the $\eta_n$ of eq. 4.6 are independent, identical distributed random variables of zero mean and variance $\sigma^2$. We thus have (the second passage follows by subtracting 8.12 from 4.7):

$$\mathbb{V}[c(v)] = E[(\sum_{k=0}^{N-1} v_k \sum_{j=0}^{k-1} \eta_j)^2] =$$

$$= E[\sum_{k=0}^{N-1} \eta_k \sum_{j=k}^{N-1} v_j] = \sigma^2 \sum_{k=0}^{N-1} (\sum_{j=k}^{N-1} v_j)^2. \tag{4.21}$$

The variance is, as expected, bilinear in the trading trajectory $v$. We can therefore express it as:

$$\mathbb{V}[C(v)] = \sum_{k,j} V_{k,j} v_kv_j, \tag{4.22}$$
where $\mathcal{V}$ is, using \eqref{eq:4.21}:
\[
\mathcal{V} = \sigma^2 \cdot \begin{pmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & 1 & 1 & \cdots & 1 \\
0 & 1 & 2 & \cdots & 2 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 1 & 2 & \cdots & (N-1)
\end{pmatrix}.
\tag{4.23}
\]

We now have all necessary tools to formulate the problem of minimizing execution costs with a risk-aversion term. Following the construction of \cite{27}, we add the variance term multiplied by a coefficient $\lambda$ (which represents our risk aversion parameter) to the transaction costs of equation \eqref{eq:4.20}. We obtain a total costs function $c_{\text{tot}}(v)$:
\[
c_{\text{tot}}(v) = v^T [I + \lambda \mathcal{V}] v + \delta 1^T |v|.
\tag{4.24}
\]

The minimum of the objective function in equation \eqref{eq:4.24} will be found using numerical methods, in Chapter \ref{ch:8}.

### 4.4.1 Solution neglecting spread costs

We note that, once again, we could express an analytical solution that minimizes the objective function \eqref{eq:4.24}, in case we neglect the spread costs. Setting $\delta = 0$, we define $\mathcal{F} \equiv I + \lambda \mathcal{V}$.

We have therefore, by the same procedure used in section \ref{sec:4.2}
\[
v^* = z \mathcal{F}^{-1} 1 = \frac{V}{1^T \mathcal{F}^{-1} 1} \mathcal{F}^{-1} 1.
\tag{4.25}
\]

In this case, instead, the solution is again analytical.
Chapter 5

Estimation methods of the propagator model for market impact

In this Chapter we develop the methods necessary to estimate empirically the impact function $f(v)$ and the market impact propagator $G_0$.

Section 5.1 is dedicated to the procedure we apply to estimate the impact function $f(v)$. Similar procedures have already been applied in other works (for example [20]), but were never explained in such detail. The adaptive binning algorithm of section 5.1.1 is original. In section 5.2 we instead explain how to estimate the propagator $G_0$. This procedure is new: we developed it, the idea arose in a discussion with J.-P. Bouchaud.

5.1 Estimation of market impact of an individual transaction

We now describe the method we use to estimate the impact function $f(v)$, which we defined in section 3.6. We have two series $r_t$ and $v_t$: the first is the series of returns over each step $r_k = p_{k+1} - p_k$. The second are the signed volumes $v_k$ traded on the market at each step, between time $t_k$ and $t_{k+1}$. The impact function $f$ was defined as the expected value of the return $r_k$ conditioned to a certain value of volume $v_k$ traded in the same $k$-th step:

$$f(v) = E[r|v].$$ (5.1)

This expected value is estimated by averaging over the series. We therefore proceed by partitioning (binning) the range of volumes, and estimating the average value of return for every bin of volumes.

In order to remove a small number of outliers we choose a cutoff volume $V_{max}$. We discard the points outside the range of volumes $[-V_{max}, V_{max}]$.

We want to partition this range, and thus we fix a step $v_{step}$ for binning such that $V_{max} = M \cdot v_{step}$ with $M \in \mathbb{N}$. The smaller we choose $v_{step}$, the more fine-grained is our estimation of the impact function (because we have more bins). However, as it will become evident in the following, if $v_{step}$ is too small the bins contain too few points and the estimation is excessively noisy.
We can now build a series $I$ of intervals of volume:

$$ I \equiv \{ [-V_{\text{max}}, (1 - M) \cdot v_{\text{step}}), ..., [-v_{\text{step}}, 0), [0, v_{\text{step}}), ..., [(M - 1) \cdot v_{\text{step}}, V_{\text{max}}) \}. \quad (5.2) $$

We have therefore partitioned the whole range under consideration. Every volume interval $i \in I$ is indexed by its midpoint $\bar{v}_i$. We call $I_v$ the series of midpoints:

$$ I_v \equiv \{ -(2N - 1), -(2N - 3), ..., -1, 1, ..., (2N - 3), (2N - 1) \} \times \frac{v_{\text{step}}}{2}. \quad (5.3) $$

We define the impact function $f(\bar{v}_i)$ on every midpoint of the class $I_v$:

$$ f : I_v \to \mathbb{R}. \quad (5.4) $$

We then consider the empirical series $v_n$ of volumes traded on the market. We have discarded the outliers outside $[-V_{\text{max}}, V_{\text{max}})$, so each $v_k$ belongs to an interval $i \in I$. The value of the function $f$ on the point $\bar{v}_i$ is computed as the average of the returns $r_k$ associated to the volumes $v_k \in i$:

$$ f(v_i) = \langle r_k | v_k \in i \rangle, \quad (5.5) $$

If there are too few volumes $v_k$ in the interval $i$ we may have large fluctuations of $f(v)$. We therefore would re-bin the sample with a larger $v_{\text{step}}$, in order to have bigger intervals $i$ that contain more empirical volumes $v_k$.

### 5.1.1 Adaptive binning algorithm

If the number of volumes $v_k$ that belong to an interval $i \in I$ is too small, the average of returns of eq. 5.5 may have too strong fluctuations (because of the noise on returns $r_k$). Therefore we need to be sure that there are at least a number $m$ volumes $v_k$ in every bin $i$. In other words, we want $\# v_k \in i \geq m$.

The standard procedure is to re-bin the sample with a larger step $v_{\text{step}}$ until there are enough volumes in each bin. We instead coded a smarter solution. Once the partition $I$ has been built, we check if there is an interval $i \in I$ with less than $m$ volumes $v_k$. If at least one of such intervals exists, we consider the interval $j \in I$ with less volumes overall. We join it to the least populated of its two neighbours ($j + 1$ or $j - 1$). We then repeat until there are no intervals $i \in I$ with less than $m$ points:

```plaintext
while \( \exists i : (\# v_k \in i) < m \) do
    \( j = \min_{j \in I} (\# v_k \in j) \)
    if \( (\# v_k \in j + 1) > (\# v_k \in j - 1) \) then
        \( j = j \cup (j - 1) \)
    else
        \( j = j \cup (j + 1) \)
end if
end while.
```
5.2 Estimation of the propagator

Once we have estimated the impact function \( f(v) \) we can consider the problem of the empirical estimation of the market impact propagator \( G_0 \). We study how to estimate \( G_0 \) from a dataset containing a series of log-returns \( r_n \) and the corresponding impacts of volumes traded on the market \( f(v_n) \). These quantities are related by eq. 3.4, which becomes:

\[
    r_n = G_0(1) f(v_n) + \sum_{k=1}^{n} [G_0(k + 1) - G_0(k)] f(v_{n-k}) + \eta_n. \tag{5.6}
\]

Our estimation method is based on linear regression and it is original. We developed the procedure, the idea arose in a discussion with J.-P. Bouchaud.

5.2.1 Estimation by linear regression

We start by defining \( g(k) \equiv G_0(k + 1) - G_0(k) \), the 1-lag difference of propagator \( G_0 \). Recalling that \( G_0(0) = 0 \), equation 5.6 becomes:

\[
    r_n = \sum_{k=0}^{n} g_k f(v_{n-k}) + \eta_n. \tag{5.7}
\]

We assume that, for large lags \( t \), the impact propagator \( G_0(t) \) converges to a value \( \theta \). It is in fact the permanent impact of the models reviewed in chapter 2. Therefore, the 1-lag difference \( g_t \) must converge to 0 and the contribution of lags \( k \) to the sum of equation 5.7 may be negligible.

We will see in Chapter 7 that a maximum lag \( k_{\text{max}} \) can be chosen such that we may truncate the sum of equation 5.7 without losing descriptive power of the model:

\[
    r_n = \sum_{k=0}^{k_{\text{max}}} g_k f(v_{n-k}) + \eta_n. \tag{5.8}
\]

This is a standard problem of linear regression. The vector of returns \( r_n \) is called dependent variable, while \( f(v_{n-k}) \) are the regressors. They both are known from the data. The vector \( g(k) \) of parameters are instead estimated by the regression procedure. It is called parameter vector or regression coefficients.

We assume that the noise term \( \eta_n \) is a random IID variable. Then it can be shown that the Ordinary Least Square (OLS) method gives the best solution to the problem of equation 5.8.

5.2.2 OLS general solution

The OLS solution to a linear regression problem consists in minimizing the Sum of Square Residuals (SSR). We define the signed impact volumes for negative times \( f(v_n) = 0 \) if \( n < 0 \). Let \( g' \) be a trial solution for equation 5.8, the SSR is given by:

\[
    S(g') = \sum_{i=0}^{L} \left( r_i - \sum_{k=0}^{k_{\text{max}}} g'(k) f(v_{i-k}) \right)^2, \tag{5.9}
\]
where $L$ is the length of the series $r_n, v_n$ of our dataset.

To simplify the expression we introduce a diagonal constant matrix $F$ such that $F_{i,k} = f(v_{i-k})$. The previous equation becomes:

$$S(g) = \sum_{i=0}^{L} \left( r_i - \sum_{k=0}^{k_{\text{max}}} g(k) F_{i,k} \right)^2 = \left( r_i - \sum_{k=0}^{k_{\text{max}}} g(k) F_{i,k} \right)^T \left( r_i - \sum_{k=0}^{k_{\text{max}}} g(k) F_{i,k} \right),$$  \hspace{1cm} (5.10)

where $T$ denotes the transpose. In vector form, calling $F = F_{i,k}$, $g = g(k)$ and $r = r_i$:

$$S(g) = (r - Fg)^T (r - Fg)$$  \hspace{1cm} (5.11)

The function $S(g)$ is quadratic in $g$ with positive-definite Hessian, and therefore it possesses a unique global minimum $\hat{g}$:

$$\hat{g} = (F^T F)^{-1} F^T r$$  \hspace{1cm} (5.12)

**Numerical solution**

The total length of the series $L$ is of the order $L \sim 5 \cdot 10^5$ for each stock we will work on. If we choose a maximum lag $k_{\text{max}} \sim 10^3$, (a typical value, as we will see in Chapter 7) the matrix $F$ has $\sim 5 \cdot 10^5$ rows and $\sim 10^3$ columns. It is too big to be stored and manipulated.

Instead, we consider the two terms composing equation 5.12: the $k_{\text{max}} \times k_{\text{max}}$ square matrix $F^T F$ and the vector $F^T r$ of length $k_{\text{max}}$. We may take advantage of the particular form of the $F$ matrix, which is diagonal constant, and compute both terms with a closed form expression (without building $F$):

$$(F^T F)_{i,k} = F^T_{i,j} F_{j,k} = \sum_{j=0}^{L} f(v_{j-i}) f(v_{j-k}),$$  \hspace{1cm} (5.13)

and

$$(F^T r)_l = F^T_{l,m} r_m = \sum_{m=0}^{L} f(v_{m-l}) r_m.$$  \hspace{1cm} (5.14)

With this definition we solve the problem of the excessive dimension of the matrix $F$, and we also make the whole procedure much faster. In addition, we will see in section 5.2.4 that the equation 5.13 and 5.14 are useful also to compute other important quantities.

**5.2.3 Errors on the estimate**

Now that we have obtained the best estimate of regression coefficients $\hat{g}$ we may be interested in knowing what is their associated error. We note that the only source of errors in the estimate is the noise terms $\eta_t$, which are not directly observed. They are estimated *a posteriori*, after having computed the regression coefficient $\hat{g}$:

$$\eta_n = r_n - \sum_{k=0}^{k_{\text{max}}} \hat{g}(k) f(v_{n-k}).$$  \hspace{1cm} (5.15)
5.2. ESTIMATION OF THE PROPAGATOR

We call $M$ the covariance matrix of the vector $\eta_n$, which corresponds to the covariance matrix of the actual observation $r$ (because the other component of $r$ is deterministic). Following standard techniques of regression analysis (for example see [37]), we may extract an equation which relates the errors on the observations $r$ with the errors on the estimated parameters $\hat{g}$. In fact, if $M$ is the covariance matrix of the vector $r$ and $M_\beta$ the covariance matrix of parameters $\hat{g}$ we have that, applying the standard theory of error propagation on eq. 5.12:

$$M_\beta = (F^T F)^{-1} F^T M F (F^T F)^{-1}. \quad (5.16)$$

we can simplify this equation by assuming that the random shocks $\eta_t$ are IID random variables. This assumption translates, in matrix form, in asserting that $M$ is a multiple of the identity matrix $I$:

$$M = \sigma^2 I. \quad (5.17)$$

In fact, independence of returns means that $M$ is diagonal, while homoscedasticity implies that all diagonal entries are equal. We therefore obtain a simple solution of the problem:

$$M_\beta = \sigma^2 (F^T F)^{-1}. \quad (5.18)$$

it is interesting to note that the errors on the estimated parameters $\hat{g}$ can be correlated even if the noise terms $\eta_t$ are not.

We want to express the errors on the elements of the time propagator $G_0(t)$, defined as $G_0(t) = \sum_{k=1}^{t} \hat{g}(k), \quad \forall t > 0$. We thus propagate the errors we just computed on the estimation of $\hat{g}(t)$. We have that:

$$\text{Var}[G_0(n)] = \text{Var}\left[\sum_{k=0}^{n} \hat{g}(n)\right] =$$

$$= \sum_{k=1}^{n} \text{Var}[\hat{g}(k)] + 2 \sum_{0<i<j<n} \text{Cov}[\hat{g}(i), \hat{g}(j)] = \sum_{i=1}^{n} \sum_{j=1}^{n} M_{\beta i,j}, \quad \forall n > 0 \quad (5.19)$$

as can be easily verified. The variance of the $n$-th term of $G_0$ is therefore the sum of entries of a $n \times n$ submatrix of $M_\beta$.

5.2.4 Goodness of the regression

The last feature of the linear regression procedure we are interested in is how well the model of equation 5.8 explains the data. We report here the linear regression equation:

$$r_n = \sum_{k=0}^{k_{\text{max}}} g_k f(v_{n-k}) + \eta_n. \quad (5.20)$$

Intuitively, we want to quantify how important is the contribution to the actual returns $r_n$ from the sum of regression coefficients $\sum_{k=0}^{k_{\text{max}}} g_k f(v_{n-k})$. The more important is that sum with respect to the random shocks $r_n$, the better our model explains the returns.

---

1 This assumption greatly simplifies the subsequent computation, but may not hold for estimations over long periods (weeks).
We therefore study the coefficient of determination $R^2$ (see for example [37]). It is defined as:

$$R^2 = 1 - \frac{\sum \eta_n^2}{\sum r_n^2}. \quad (5.21)$$

We can think of $R^2$ as the fraction of volatility explained by the model. In fact, $\sum \eta_n^2 / \sum r_n^2$ is the ratio between unexplained variance (variance of the model’s errors $\eta_n$) and total variance of the data. The coefficient ranges in $[0, 1]$:

- if $R^2 = 0$ it follows that $\eta_n = r_n$. Therefore, the linear regression model of eq. 5.20 explains no component of returns $r_n$. The only contribution to the returns comes from the random shocks $\eta_n$. Therefore, we say that model explains no volatility;

- if $R^2 = 1$ then $\eta_n = 0$. The linear regression model explains completely the returns $r_n$. There is no contribution from the random noise. The model explains all the volatility.

We now express the coefficient $R^2$, recalling equation 5.11 for the sum of squared residuals:

$$\sum \eta_n^2 = S(g) = (r - Fg)'(r - Fg) = r'r - 2g'(F'r) + g'(F'F)g. \quad (5.22)$$

We obtain:

$$R^2 = 1 - \frac{S(g)}{r'r} = \frac{2g'(F'r) - g'(F'F)g}{r'r}. \quad (5.23)$$

We recall that $F'r$ and $F'F$ were already computed for the actual regression, as explained at the end of section 5.2.2. Therefore, we can compute the $R^2$ coefficient with little numerical effort.

All the methods for the $G_0$ estimation we explained in this Chapter will be applied on actual financial data in Chapter 7.
Chapter 6

Data and methods

This Chapter describes the data and methods of the empirical studies of this work. In section [6.1] we make a brief description of the financial dataset we use in the empirical estimates of the two following Chapters. We also explain some of the filtering methods we have applied in order to perform the subsequent computations. Section [6.2] focuses instead on the software and libraries we have used to code and run our empirical estimations.

6.1 Datasets description

For our empirical research we have used two different datasets reporting the trading activity in the London Stock Exchange (LSE). The first one covers from May, 2000 to December, 2002. The second one is much more recent, from January 2011 to April 2011. The latter was provided by the Linear Quantitative Research Group of J. P. Morgan. Both datasets contain data for many different stocks (92 the first and 100 the second). For simplicity we decided to work on only a selection of them, namely two stocks for each dataset.

6.1.1 2000 - 2002 dataset

These data come from the LSE’s electronic order book Stock Exchange Trading System (SETS). It includes details of all order book event (submissions of limit and market orders, cancellations of limit orders) on 92 among the most liquid stocks traded. It spans more than 2 years: from 1st May, 2000 to 31st December, 2002. SETS was at that time (and still is) one of the most mature electronic order book markets, with high liquidity and a dominant proportion of all traded volume occurring on the electronic order book rather than through competing mechanisms (such as over the counter transactions, private deals between big investors). This dataset was filtered for an old work [38], obtaining a sequence of all the trades on the market. For each trade there is therefore a line containing the time of the trade, its volume size, price, the best bid and best ask before and after the trade. Also, we know whether the transaction was initiated by the buyer or the seller, because in the original
dataset we have a log of all market orders. Figure 6.1 shows a sample of the filtered dataset, and explains the variables in every column.

Figure 6.1: Sample of the filtered 2000 - 2002 dataset, for the stock AZN. The first column is the date (2 May 2000). The second is the trade direction: 1 (2) means that it is buyer (seller) initiated. The third and fourth are the volume traded and transaction price. Fifth to eighth columns are the best bid and ask prices before the transaction, and the best bid and ask after the transaction. The ninth columns is not used. The last column is the time.

We have to focus on only a selection of all the 92 stocks. We decided to select only two of them. In order to choose them, we consider an empirical fact of the stock market. Prices are not allowed to vary continuously: they have a finite step (tick size) fixed by the exchange regulators. It can range from 1 penny to 1 pound or more, depending on the price of the stock, for example when the stock price is between 10 and 50 pounds the tick size is fixed at 5 pennies.

As we will see, the tick size - price ratio can be of the same order of magnitude as the market impact. It may therefore play an important role.

In order to test our model against the most varied conditions, we chose the two stocks with the most different average price - tick size ratio. We therefore chose the stock with the highest price - tick size ratio (Astrazeneca, symbol AZN), and the one with the lowest (Vodafone, symbol VOD). In table 6.1 their properties are summarized.

6.1.2 2011 January - April dataset

The second is a trade and quote dataset of the London Stock Exchange by Thomson-Reuters (a financial information company). A trade and quote dataset contains two kind of lines:

- trades, the series of all transactions with its time, volume and the price;
- quotes, updates on the current best bid and ask in the Limit Order Book. They contain the current best bid and ask prices, and the volumes which are offered at those prices (sum of the volumes of all limit orders at those prices).

It is considerably more recent, it was provided by the Linear Quantitative Research Group of J. P. Morgan in London. It comprises the 100 most capitalized stocks on the LSE,
6.1. DATASETS DESCRIPTION

which together compose the FTSE100 stock index. In figure 6.2 we can see a sample of the dataset, on the stock BT.

Figure 6.2: Sample of Thomson-Reuters trade and quote 2011 dataset, for the stock BT. The first column is the stock symbol. The second, third and fourth are the date, time, and offset with respect to London time. The fifth column tells whether the line is a trade or a quote. If it is a trade we then have the trade price and volume. If it is a quote we then have the best bid, size offered at best bid, best ask and size offered at best ask.

In order to use these data for our estimation, we transformed it in a form similar to the filtered dataset of figure 6.1, a series of all trades. We therefore select all trade lines of the dataset of figure 6.2 For each of them:

- we search if there is another neighboring trade line at the same time and same price. If there is, we merge the two lines, adding the volume size of the trades. in fact, we found that sometimes a single trade is split in more then one line;
- we select the last quote line before the trade line under consideration. We therefore know the best bid and best ask just before the trade;
- we infer the trade direction, i.e. whether the transaction was initiated by the buyer or by the seller. We apply the Lee & Ready algorithm (see [39]). We observe the price $\tilde{P}_n$ of the trade. We also know the midprice $P_n$ just before the trade (average of bid and ask). If the trade price is bigger that the midprice, $\tilde{P}_n > P_n$, then we classify the trade as buyer initiated. If instead the price is smaller that the midprice $\tilde{P}_n < P_n$, then it is seller initiated. The case $\tilde{P}_n = P_n$ remains undetermined, but it represents a very small fraction of the total (< 0.1%).

We then copy every line in a new file, and obtain what is shown in figure 6.3. We will use this filtered version of the data in the subsequent computations.

As for the other dataset, we have to choose a small selection of stocks to be used in the rest of the work. We applied the same criterion as before, analyzing the average tick size - price ratio. We chose the stocks Barclays (symbol BARC) and British Petrol (symbol BP) which have respectively a very high and low tick size - price ratio. The properties of the various stocks are shown in table 6.1.
Figure 6.3: Sample of the our filtered version of the dataset of figure [6.2] for the stock BT. We have the series of all trades. The first and second columns are the date and time. The third and fourth are trade volume and price. The fifth and sixth columns are the best bid and ask just before the trade. The last column is the trade direction (+1 if initiated by the buyer, -1 by the seller).

Table 6.1: The variables we have described for the four stocks we consider. The number of daily transaction refers to the number of distinct trades that take place in a market day. The average tick-size price ratio is expressed in basis points (bp), which equals $10^{-4}$, a customary quantity in finance: in other words, 3.7 (bp) = $3.7 \cdot 10^{-4}$.
6.2 Software and libraries used

This work has required a great effort of development and coding. There were four main tasks:

- we first worked on data processing and filtering. We implemented the procedure we described in section 6.1.2 to process the 2011 dataset, in order to obtain what is shown in figure 6.3;
- then we coded the estimation procedures of the impact function $f(v)$ and propagator $G_0$, as described in Chapter 5;
- the biggest computational challenge was to find the optimal trading schedules in all different settings described in Chapter 4. In fact, in most of the cases we used complicated numerical optimization methods. We will give more details when we show the results in Chapter 8;
- at last, we printed all the plots presented in this Thesis (unless stated otherwise in the caption). Most of these plots are in Chapters 7 and 8.

We wrote almost all our code in the Python programming language (http://www.python.org), using mainly the scientific library SciPy [40] (in particular the numerical extension NumPy). We made all the graphs with the python plotting library Matplotlib [41]. The procedure for the filtering of the 2011 dataset was instead coded in C++ [42], because of the great amount of data to input and output.

We chose Python because of its great flexibility, easy readability of the code and fast development process. Also, it is free and open-source. With the addition of the scientific libraries we mentioned we had all the data structures and functions we needed.
Chapter 7

Empirical estimate of propagator model

This chapter is devoted to the empirical estimate of the parameters of the impact propagator model, with the methods that we developed in Chapter 5. We will work on the data we described and analyzed in Chapter 6. All material in this Chapter is original. We will use the model of market impact propagator in three different settings. They differ by the definition of the series of times \( t_n \), as discussed in section 7.1. Section 7.2 is devoted to the analysis in trade time. In section 7.3 we instead use time aggregation by a constant number of transactions. Section 7.4 uses the aggregation model of constant real time intervals. We will use the results of this section to calibrate the optimal execution strategies in Chapter 8. Section 7.5 compares the three different approaches by studying how well the propagator model fits the data in different time settings.

7.1 Different times for financial markets

In Chapter 3 we have reviewed the propagator model, which has arisen in the econophysics literature. In this Chapter we are going to estimate this propagator empirically, working on the datasets we have studied in Chapter 6. We will use three different definitions of time series, trade time, aggregated trade time and real time. They will lead to quite different results.

7.1.1 Trade time

The propagator formalism was originally proposed in trade time (see [18], [22]). It is a discrete time setting in which each transaction on the market increases time by one unit. We fix \( t^t_0 \) as the time of the first transaction of the period we consider. The superscript “\( tt \)” stands for trade time. Then, each transaction increases \( n \) by one unit and thus we have that \( t^t_n \) is the time of the \( n \)-th transaction.

The price series is easily defined: \( p^t_n \) is the log-midprice of the stock right before the \( n \)-th transaction. We also recall the returns definition: \( r^t_n = p^t_{n+1} - p^t_n \).
CHAPTER 7. EMPIRICAL ESTIMATE OF PROPAGATOR MODEL

The volume \( v_n \) is simply the volume in shares of the \( n \)-th transaction. It is positive if the transaction was buyer initiated and negative if seller initiated. It is the easiest setting to use for the estimate of the propagator \( G_0 \). We will apply it in section 7.2.

### 7.1.2 Aggregated trade time

This definition of time is a generalization of trade time. In trade time we increase time by one unit after each transaction. In aggregated trade time we aggregate \( d \in \mathbb{N} \) transactions, so that time is increased by one unit after \( d \) transactions on the market. We fix \( t_0^{att} \) as the time of the first transaction of the period we consider. The superscript “att” stands for aggregated trade time. Then, every \( d \) transactions we increase time by one unit and we have that \( t_n^{att} \) is the time of the \((n \cdot d)\)-th transaction.

The security price \( p_n^{att} \) is defined as the log-midprice right before the \((n \cdot d)\)-th transaction. The volume \( v_n^{att} \) is the algebraic sum of volumes of the \( d \) transactions between time \( t_n^{att} \) and time \( t_{n+1}^{att} \). We can express the volume \( v_n^{att} \) in term of trade time volumes \( v_n^{tt} \):

\[
v_n^{att} = \sum_{i=d \cdot n}^{d \cdot n + (n+1)-1} v_i^{tt}.
\]

(7.1)

Where each trade time volume \( v_i^{tt} \) is positive if the transaction was buyer initiated and negative if seller initiated.

If we fix the aggregation number \( d = 1 \), we obtain again the trade time definition. Thus, aggregated trade time is in fact just a generalization of trade time. However, the impact propagator \( G_0 \) has never been studied in this setting. We will do so in section 7.3.

### 7.1.3 Real time

The last definition of time we consider is probably the most natural one. We fix \( t_0 \) as the opening time of the market on the period we consider. We then define the series of times \( t_n \) as separated by a constant time interval, which will be 5 minutes in the rest of this work. For example, we could have \( t_0 = 8:00 \), \( t_1 = 8:05 \), \( t_2 = 8:10 \),...

The stock price \( p_n \) is defined as the log-midprice right before time \( t_n \). This definitions have already been used in Chapter 2 when reviewing the models for optimal execution of [26], [27] and [28].

The series of volumes \( v_n \) is now defined as the algebraic sum of all volumes \( v_i^{tt} \) traded on the market between time \( t_n \) and \( t_{n+1} \):

\[
v_n = \sum_{t_i^{tt} \in [t_n, t_{n+1}]} v_i^{tt}.
\]

(7.2)

By this notation we mean that we sum over all volumes \( v_i^{tt} \) of transactions whose times \( t_i^{tt} \) are contained in the interval \([t_n, t_{n+1}]\).

This is what is called volume imbalance over a certain time interval. This quantity may be very big in absolute value during time intervals of high market activity, like the ones near
market opening and closing. On the other side, its absolute value can be small during intervals when there is not much trading. Therefore, we will concentrate on another related variable, the normalized volume imbalance $v_n^{nor}$. It is obtained by dividing the volume imbalance $v_n$ in the $n$-th interval by the total volume traded in the interval, i.e. the sum of absolute values of the volumes $v^t_i$ traded:

$$v_n^{nor} = \frac{\sum_{t^t_i \in [t_n,t_{n+1}]} v^t_i}{\sum_{t^t_i \in [t_n,t_{n+1}]} |v^t_i|}$$  \hspace{1cm} (7.3)

If there is no trade in the $n$-th interval we define $v_n^{nor} = 0$. We note that every $v_n^{nor}$ is bounded by $v_n^{nor} \in [-1, 1]$, a feature which will prove very useful in the following. This definition is similar to the one found in [1], section Order flow in real time. In section 7.4 we apply it to estimate the market impact propagator $G_0$.

We can now turn to the actual estimations, starting with trade time.

### 7.2 Model in trade time

We start our estimations of market impact propagator in the framework of trade time $t^t_n$, which was defined in section 7.1.1: each transaction on the market increases time by one unit. It is the simplest setting to attempt a functional estimation and fit of the propagator $G_0$: in fact some results of similar estimations have already been published (see [22]).

As we explained in Chapter 5, we first need to estimate the impact function $f(v^t)$, which represents the volume dependence of market impact. Then we will concentrate on the actual propagator $G_0$.

#### 7.2.1 Volume dependence of market impact

The first step in estimating the impact propagator is to model the volume dependence of market impact. It is the impact function $f$ we defined in section 3.2.1.

In trade time, this corresponds to the market impact of individual transaction we discussed in Chapter 1, section 1.5.2. In fact this is a well known problem already examined in many studies (see [19] for a review).

We recall that the impact function $f(v^t)$ is the expected value of return $r^t$, conditioned to a value $v^t$ of volume exchanged:

$$f(v^t) = E[r^t|v^t].$$  \hspace{1cm} (7.4)

This can be estimated on the data following the procedure we explained in section 5.1 on each of the four stocks we choose the cutoff volume $V_{max}$, the step for binning $v_{step}$ and the minimum number of points per bin $m$ (for the adaptive binning algorithm described in section 5.1.1).

We fix $V_{max}$ in order to remove a small number of big volume outliers. In practice, we choose $V_{max}$ such that the number of elements $v^t_n \in [-V_{max},V_{max}]$ represents more than 99.9% of the total sample for each stock. Instead, $v_{step}$ and $m$ are chosen such that the plots we obtain have an acceptable level of noise.
CHAPTER 7. EMPIRICAL ESTIMATE OF PROPAGATOR MODEL

Figure 7.1 shows the result of the estimations. In the caption we can see the parameters chosen for each stock.

We then attempt a functional fit of the impact function $f(v^t)$. The function that has proven most successful in the literature is a power of the absolute value of the volume $|v^t|$ multiplied by the sign of the volume:

$$f(v^t) = \theta \cdot \text{sign}(v^t) \cdot |v^t|^\phi.$$  \hspace{1cm} (7.5)

We can see in figure 7.1 that it fits well the data, and the exponent $\phi$ of the powers for the four stocks are all small positive reals $\phi \in [0.01, 0.5]$. The results of the fit are shown in the caption of the figure, alongside the errors provided by the fitting procedure.

The two stocks of the 2011 dataset, BARC and BP, have considerably smaller exponents then the stocks of the 2000-2002 dataset, VOD and AZN. In fact, their impact function $f$ looks very similar to a step function. Therefore, we may think that some market mechanisms have changed in these last 10 years, a feature which we will further investigate in this Chapter.

7.2.2 Propagator

We can now turn to the estimation of the market impact propagator $G_0$, following the procedure described in section 5.2. We run a linear regression of the returns $r^t$ on the series of impact function $f(v^t)$ of the past volumes transacted, as in equation 5.8. Recalling that $g(k) \equiv G_0(k+1) - G_0(k)$, the regression model is:

$$r^t_n = \sum_{k=0}^{k_{\text{max}}} g_k f(v^t_{n-k}) + \eta_n.$$  \hspace{1cm} (7.6)

For each of the four stocks we fix the maximum lag $k_{\text{max}}$ for the regression of the propagator. As we discussed earlier, $k_{\text{max}}$ should be chosen such that the contribution to the sum of the terms $(g_k f(v^t_{n-k}))$ for $k > k_{\text{max}}$ is not significant.

In order to define which are the $k$ lags for which the contribution of $g_k$ to the model is significant, we recall that we developed (in section 5.2.4) a procedure to estimate the coefficient of determination $R^2$ of the linear regression model. This coefficient $R^2$ measures how well the model fits the data. The bigger it is, the better.

We therefore chose to fix the maximum lag $k_{\text{max}}$ to a value such that extending the regression to $k > k_{\text{max}}$ does not improve the coefficient of determination by more than 1%. Following this criterion, we fixed $k_{\text{max}} = 1000$ lags for the two stocks of the 2000-2002 dataset, and $k_{\text{max}} = 2000$ for the two stocks of the 2011 dataset.

Therefore, we obtained from the linear regression the parameters $g(k)$. From them, we can extract the propagator $G_0$, recalling that $G_0(0) = 0$. In fact we have that:

$$G_0(k + 1) = G_0(k) + g(k).$$  \hspace{1cm} (7.7)

In figure 7.2 we show the propagator $G_0$ we obtained for the four stocks considered.

As we discussed in section 5.2.3 we have a procedure to know the errors on the parameters $g(k)$ of the linear regression. From those, we can derive the variance, and thus the standard deviation $\sigma$, on each element $G_0(k)$ of the propagator.
Figure 7.1: Estimation of the impact of individual transactions, on the stocks we consider. We plot on the $x$ axis the transaction volume, and on the $y$ the corresponding average return. The parameters chosen for the estimation procedure are: $V_{\text{max}} = 30 \cdot 10^4$, $v_{\text{step}} = 200$, $m = 120$ for VOD, $V_{\text{max}} = 10 \cdot 10^4$, $v_{\text{step}} = 300$, $m = 70$ for AZN, $V_{\text{max}} = 10 \cdot 10^4$, $v_{\text{step}} = 100$, $m = 120$ for BARC, $V_{\text{max}} = 10 \cdot 10^4$, $v_{\text{step}} = 100$, $m = 120$ for BP. The parameters obtained from the fit of equation 7.5 are: $\theta = (0.20 \pm 0.01)$ bp, $\phi = 0.313 \pm 0.005$ for VOD, $\theta = (0.83 \pm 0.08)$ bp, $\phi = 0.195 \pm 0.008$ for AZN, $\theta = (0.55 \pm 0.04)$ bp, $\phi = 0.014 \pm 0.009$ for BARC and $\theta = (0.31 \pm 0.03)$ bp, $\phi = 0.04 \pm 0.01$ for BP.
CHAPTER 7. EMPIRICAL ESTIMATE OF PROPAGATOR MODEL

Figure 7.2: Estimation of the impact propagator $G_0$ on four stocks we consider, in trade time. The $R^2$ coefficients of the linear regression are 0.224 for AZN, 0.257 for VOD, 0.083 for BP and 0.090 for BARC. The fit parameters are $\Gamma_0 = 1.102 \pm 0.002$, $l_0 = 7.60 \pm 0.09$, $\beta = 0.1729 \pm 0.0005$ for VOD, $\Gamma_0 = 1.33 \pm 0.002$, $l_0 = 7.28 \pm 0.07$, $\beta = 0.1640 \pm 0.0004$ for AZN, $\Gamma_0 = 2.23 \pm 0.01$, $l_0 = 16.6 \pm 0.2$, $\beta = 0.271 \pm 0.001$ for BARC, and $\Gamma_0 = 1.95 \pm 0.01$, $l_0 = 14.7 \pm 0.2$, $\beta = 0.246 \pm 0.001$ for BP.
7.3 Model in aggregated trade time

In the figure we plot the propagator $G_0(k)$ and its error bands of one standard deviation $\sigma$. In the caption we also state the values of the $R^2$ coefficients of the regressions.

From the discussion on long memory of the trade sign series in Chapter 1 (section 1.5.3) we expect a decreasing impact propagator with the lag $k$, which is indeed observed.

We also attempted to fit a functional form of the propagator. The function $G_0^{fit}$ which was proposed in the literature, in particular [22], is a decaying power law with a correction for small lags $l$:

$$G_0^{fit}(l) = \frac{\Gamma_0}{(l_0^2 + l^2)^{\beta/2}},$$

(7.8)

where $\Gamma_0$ is a multiplicative constant, $l_0$ is a correction for small lags and $\beta$ the coefficient of the power law decay. In fact, we note that for $l \gg l_0$:

$$G_0^{fit}(l) \approx \frac{\Gamma_0}{l^\beta}.$$  

(7.9)

This function fits quite well the data, as we can see in figure 7.2. The parameters obtained from the fit are shown in the caption as well as the errors provided by the fitting routine.

7.3 Model in aggregated trade time

We now turn to the setting of aggregated trade time, defined in section 7.1.2. We fix an integer $d$, such that we increase time by one unit every $d$ transactions on the market.

We have tried the model with two different aggregation numbers, $d = 8$ and $d = 64$, so that we can better observe the effects of aggregation on the propagator model.

The impact function $f(v^{att})$ in the setting of aggregated trade time has already been estimated in some past works [19]. However, the propagator has never been studied empirically in aggregated trade time.

7.3.1 Volume dependence of market impact

We start by estimating the impact function $f(v^{att})$, which has already been studied in [19], section 5.2. In that work it is found that the impact function $f(v^{att})$ has a linear section for small values of $v^{att}$.

We show in figure 7.3 the plots of impact function in aggregated trade time with number of aggregation $d = 8$. In figure 7.4 we instead show the plots with aggregation $d = 64$. In the captions we provide the parameters $V_{max}$, $v_{step}$ and $m$ chosen for the estimation. We followed the same criteria of section 7.2.1.

We then attempted a functional fit, with a new functional form. By observing the plots we note that the function must be linear for small volumes $v^{att}$, and reach a constant value for high values of $v^{att}$. We therefore used an arctangent function:

$$f(v^{att}) = \theta \arctan(\rho v^{att}).$$

(7.10)

We can see from the figures 7.3 and 7.4 that it fits well the data. The fit parameters are provided in the captions.
Figure 7.3: Estimation of the impact of 8 aggregated transactions, on the stocks we consider. The parameters chosen for the estimation procedure are: $V_{\text{max}} = 200 \cdot 10^4$, $v_{\text{step}} = 500$, $m = 100$ for VOD, $V_{\text{max}} = 8 \cdot 10^4$, $v_{\text{step}} = 100$, $m = 70$ for AZN, $V_{\text{max}} = 6 \cdot 10^4$, $v_{\text{step}} = 150$, $m = 60$ for BARC, $V_{\text{max}} = 6 \cdot 10^4$, $v_{\text{step}} = 200$, $m = 70$ for BP. The fit parameters for the function of eq. 7.10 are $\theta = (17.1 \pm 0.2) \text{ bp}$, $\rho = (0.0176 \pm 0.0004) \cdot 10^{-4}$ for VOD, $\theta = (12.2 \pm 0.2) \text{ bp}$, $\rho = (0.76 \pm 0.02) \cdot 10^{-4}$ for AZN, $\theta = (1.45 \pm 0.04) \text{ bp}$, $\rho = (1.3 \pm 0.1) \cdot 10^{-4}$ for BARC and $\theta = (0.99 \pm 0.03) \text{ bp}$, $\rho = (1.4 \pm 0.1) \cdot 10^{-4}$ for BP.
Figure 7.4: Estimation of the impact of 64 aggregated transactions, on the stocks we consider. The parameters chosen for the estimation procedure are: $V_{\text{max}} = 800 \cdot 10^4$, $v_{\text{step}} = 20000$, $m = 100$ for VOD, $V_{\text{max}} = 20 \cdot 10^4$, $v_{\text{step}} = 500$, $m = 70$ for AZN, $V_{\text{max}} = 30 \cdot 10^4$, $v_{\text{step}} = 2500$, $m = 60$ for BARC, $V_{\text{max}} = 30 \cdot 10^4$, $v_{\text{step}} = 2500$, $m = 70$ for BP. The fit parameters for the function of eq. 7.10 are: $\theta = (36.3 \pm 0.9)$ bp, $\rho = (0.0056 \pm 0.0003) \cdot 10^{-4}$ for VOD, $\theta = (28 \pm 1)$ bp, $\rho = (0.18 \pm 0.02) \cdot 10^{-4}$ for AZN, $\theta = (2.8 \pm 0.3)$ bp, $(0.4 \pm 0.1) \cdot 10^{-4}$ for BARC and $\theta = (2.5 \pm 0.3)$ bp, $\rho = (0.3 \pm 0.1) \cdot 10^{-4}$ for BP.
CHAPTER 7. EMPIRICAL ESTIMATE OF PROPAGATOR MODEL

7.3.2 Propagator

We can now estimate the impact propagator $G_0$ in aggregated trade time, by using the fitted function $f(v^{att})$ of eq. (7.10) in the linear regression procedure.

We choose the maximum lags $k_{max}$ by analogy with the estimation we run in trade time (section 7.2). Therefore, we divide the $k_{max}$ used in trade time setting by the aggregation number $d$.

For the aggregation $d = 8$ we obtain $k_{max} = 125$ for VOD and AZN and $k_{max} = 375$ for BP and BARC. Instead, for the aggregation $d = 64$ we have $k_{max} = 16$ for VOD and AZN and $k_{max} = 47$ for BP and BARC.

The results of the estimation procedure are shown in figures 7.5, for $d = 8$, and 7.6 for $d = 64$. We obtain the errors of the regression as explained in section 5.2.3. We thus plot the propagator and its error bands of one standard deviation. In the captions of the figures we also provide the coefficients of determination $R^2$ of the estimations.

We also fitted a functional form for the impact propagator $G_0$. We used the same function of eq. (7.8) for the trade time setting. The parameters of the fit are reported in the captions of the figures.

We note that there is no theoretical reason why the propagator $G_0$ should preserve its functional shape under aggregation.

In the figures we can also see the errors associated to the fit (as reported by the fitting routine).

7.4 Model in real time

The last version of the propagator model we study is in the setting of real time intervals, as defined in section 7.1.3. In particular, we use 5 minutes intervals.

The estimation of the impact propagator in real time has never been attempted before. We will use the parameters we obtain here to calibrate the optimal execution schedules we study in Chapter 8.

We start by studying the impact function of the normalized volume imbalances $f(v^{nor})$ we defined before (in section 7.1.3).

7.4.1 Volume dependence of market impact

We repeat once again the procedure described in section 5.1 for the estimation of the impact function. We want to estimate the impact function $f(v^{nor})$ of the normalized volume imbalance. It is defined as the expected return $r_n = p_{n+1} - p_n$ conditioned to a normalized volume imbalance $v_n^{nor}$ on the market:

$$f(v^{nor}) = E[r_n|v_n^{nor}]. \quad (7.11)$$

We recall that each $v_n^{nor} \in [-1, 1]$, because of the way we defined it. Therefore, the cutoff volume $V_{max}$ for the estimation is easily fixed at 1. We also choose, for all the four stocks, $v_{step} = 0.002$ and $m = 15$. These values provide a reasonable balance between noise suppression and readability of the plot.
Figure 7.5: Estimation of the impact propagator $G_0$ on the four stocks we consider, in aggregated trade time. The $R^2$ coefficients of the linear regression are 0.33 for VOD, 0.25 for AZN, 0.07 for BARC and 0.07 for BP. The fit parameters are $\Gamma_0 = 0.925 \pm 0.003$, $l_0 = 0.0 \pm 0.7$, $\beta = 0.083 \pm 0.001$ for VOD, $\Gamma_0 = 1.098 \pm 0.006$, $l_0 = 1.59 \pm 0.09$, $\beta = 0.156 \pm 0.001$ for AZN, $\Gamma_0 = 1.17 \pm 0.01$, $l_0 = 2.2 \pm 0.2$, $\beta = 0.184 \pm 0.002$ for BARC and $\Gamma_0 = 1.34 \pm 0.03$, $l_0 = 3.4 \pm 0.3$, $\beta = 0.221 \pm 0.005$ for BP.
Figure 7.6: Estimation of the impact propagator $G_0$ on the four stocks we consider, in aggregated trade time. The $R^2$ coefficients of the linear regression are 0.44 for VOD, 0.30 for AZN, 0.04 for BARC and 0.05 for BP. The fit parameters are $\Gamma_0 = 0.92 \pm 0.05$, $l_0 = 0.0 \pm 0.1$, $\beta = 0.106 \pm 0.004$ for VOD, $\Gamma_0 = 1.05 \pm 0.06$, $l_0 = 0.0 \pm 0.1$, $\beta = 0.17 \pm 0.01$ for AZN, $\Gamma_0 = 0.95 \pm 0.04$, $l_0 = 0.0 \pm 0.1$, $\beta = 0.189 \pm 0.009$ for BARC and $\Gamma_0 = 0.89 \pm 0.05$, $l_0 = 0.0 \pm 0.1$, $\beta = 0.19 \pm 0.01$ for BP.
In figure 7.7 we show the result of the estimation, fitting the data with an arctangent function $f_{\text{atan}}$:

$$f_{\text{atan}}(v^{\text{nor}}) = \theta \arctan(\rho v^{\text{nor}}).$$

(7.12)

The parameters resulting from the fit are shown in the caption of the figure. As we can see, the impact is very close to be linear. In Chapter 8, when we will use the impact function of the real time model to calibrate the optimal execution strategies, it will prove very useful to have a linear $f$. We thus approximate the impact function in real time with:

$$f(v^{\text{nor}}) = \theta v^{\text{nor}}.$$  

(7.13)

In figure 7.8 we can see the fit with the linear impact function.

We note that these plots are similar to the ones found in [43], where a slightly different volume definition is used.

### 7.4.2 Propagator

We now estimate the impact propagator $G_0$, using the linear function $f_{\text{lin}}(v^{\text{nor}})$ we obtained in the previous section. We have to choose a maximum lag for the linear regression $k_{\text{max}}$. We decided to fix $k_{\text{max}} = 102$, which corresponds to the number of 5 minutes intervals in a market day (from 8:00 to 16:30). Recalling the criterion we used in section 7.2.2, we tested that extending the regression to $k > k_{\text{max}}$ does not improve the coefficient of determination by more than 1%.

We obtain the impact propagators shown in figure 7.9. We plot the propagator and its error bands, with the standard deviation obtained from the procedure. We see that the estimation on the two new stocks (BP and BARC) is considerably noisier than the old ones (AZN and VOD). This could be due to the fact that the new dataset contains much less days that the old one, as we can see in table 6.1.

The fit is performed once again with the functional form of eq. 7.8. It works remarkably well with the data of the 2000-2002 dataset, but on the two stocks of the 2011 dataset it seems poor. We assume that this difference is due to noise. In fact, the fitted form of the propagator lies within the error bars of the estimation.

We can see in the caption of the figures the parameters obtained from the fit. We will use impact propagators with these parameters in the following Chapter, when we propose optimal trading schedules to minimize market impact costs.

### 7.5 Variance explained by the model

We recall the coefficient of determination $R^2$ of the linear regression. It was discussed in section 5.2.4, explaining how it is estimated. In particular, we recall its interpretation as fraction of explained variance. The coefficient $R^2 \in [0, 1]$ is the fraction of the total variance of the stock that is explained by the model. The closer it is to 1, the better our impact model describes the dynamics of prices.
Figure 7.7: Estimation of the impact in real time, on stocks we consider. We plot on the $x$ axis the normalized volume imbalance, and on the $y$ the corresponding average return (in basis points). The fit parameters of eq. (7.12) are $\theta = (20.8 \pm 0.6)$ bp and $\phi = 1.79 \pm 0.08$ for VOD, $\theta = (12.6 \pm 0.4)$ bp and $\phi = 1.72 \pm 0.06$ for AZN, $\theta = (9.9 \pm 0.5)$ bp and $\phi = 1.41 \pm 0.05$ for BARC, and $\theta = (3.0 \pm 0.2)$ bp and $\phi = 3.2 \pm 0.2$ for BP.
Figure 7.8: Estimation of the impact in real time, on stocks we consider. We plot on the $x$ axis the normalized volume imbalance, and on the $y$ the corresponding average return (in basis points). The fit parameters of eq. 7.13 are $\theta = (26.0 \pm 0.8)$ bp for VOD, $\theta = (15.4 \pm 0.5)$ bp for AZN, $\theta = (12.2 \pm 0.5)$ bp for BARC, and $\theta = (7.9 \pm 0.3)$ bp for BP.
CHAPTER 7. EMPIRICAL ESTIMATE OF PROPAGATOR MODEL

Figure 7.9: Estimation of the impact propagator $G_0$, on two stocks of the 2011 January-April dataset, in real time. The $R^2$ coefficients of the linear regression are 0.20 for AZN, 0.29 for VOD, 0.04 for BARC and 0.04 for BP. The fit parameters are $\Gamma_0 = 1.07 \pm 0.01$, $l_0 = 4 \pm 1$, $\beta = 0.075 \pm 0.002$ for VOD, $\Gamma_0 = 1.40 \pm 0.04$, $l_0 = 20 \pm 1$, $\beta = 0.190 \pm 0.006$ for AZN, $\Gamma_0 = 1.4 \pm 0.1$, $l_0 = 0 \pm 1$, $\beta = 0.18 \pm 0.02$ for BARC and $\Gamma_0 = 1.6 \pm 0.1$, $l_0 = 0 \pm 2$, $\beta = 0.21 \pm 0.03$ for BP.
7.5. VARIANCE EXPLAINED BY THE MODEL

We summarize in table 7.1 the coefficients $R^2$ for all the models studied. We also add the coefficient estimated for a real time model with 1 minute intervals that is not shown here.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>trade time transactions</th>
<th>8 aggregated transactions</th>
<th>64 aggregated transactions</th>
<th>1 minute aggregation</th>
<th>5 minutes aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZN</td>
<td>0.224</td>
<td>0.245</td>
<td>0.304</td>
<td>0.192</td>
<td>0.204</td>
</tr>
<tr>
<td>VOD</td>
<td>0.257</td>
<td>0.329</td>
<td>0.440</td>
<td>0.209</td>
<td>0.292</td>
</tr>
<tr>
<td>BP</td>
<td>0.083</td>
<td>0.070</td>
<td>0.047</td>
<td>0.054</td>
<td>0.039</td>
</tr>
<tr>
<td>BARC</td>
<td>0.090</td>
<td>0.069</td>
<td>0.044</td>
<td>0.070</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Table 7.1: $R^2 \in [0, 1]$ coefficients of the linear regression. Defined in section 5.2.4.

We can see that there are substantial differences between the $R^2$ parameters of the propagator models for the two datasets, and for the different time definitions. In particular:

- the two stocks of the 2000-2002 dataset, AZN and VOD, have in general higher parameters $R^2$ then the two stocks of the new dataset, BP and BARC. This could mean that in the last 10 years the market mechanisms have evolved in such a way that the market impact has become less important in determining the stock returns. In other words, the contribution to the returns from the random, unexplained shocks $\eta$ has become more important;

- the transition from trade time to 8 and 64 aggregated transactions has different effects. For the two stocks of the 2000-2002 dataset (AZN and VOD), the aggregation improves the model. It is impressive to note that for VOD, in the 64 aggregated transactions model, 44% of the total volatility is explained by the model. The two stocks of the 2011 dataset (BP and BARC) show instead the opposite effect: the aggregation leads to diminished coefficients of determination;

- the real time model leads to coefficients of determination of the same order of magnitude than the trade time model. This feature alone is of great appeal, because the propagator formalism has never been used before to estimate the impact in real time. Again, the stocks of the new and old dataset behave differently: 1 minute aggregation is better the 5 minutes for BP and BARC, but is worse for AZN and VOD.

In the next Chapter we propose optimal trading schedules to minimize impact costs, using the propagator model in real time.
CHAPTER 7. EMPIRICAL ESTIMATE OF PROPAGATOR MODEL
Chapter 8

Empirical calibration of optimal executions with the propagator model

In this Chapter we apply the theory we developed in Chapter 4 to the problem of optimal executions with the market impact propagator $G_0$. All the material proposed here is original.

We start in section 8.1 by recalling and specializing some definitions. Section 8.2 deals with the simplest specification of the problem, i.e. we neglect bid-ask spread costs and only minimize impact costs. We apply the theory developed in section 4.2. Section 8.3 instead adds the contribution of the cost of a constant bid-ask spread. It follows the model of section 4.3. Section 8.4 introduces risk aversion and builds the efficient frontier of optimal execution, as discussed in 2.5, and in 4.4.

8.1 Statement of the problem

We study an execution strategy with a time horizon $T$ of one market day. We model the problem of buying a quantity $X$ of shares, which we fix at 1% of the total volume traded in the market during the day we consider.

The London Stock Exchange opens at 8:00 and closes at 16:30. We choose to work with 5-minutes intervals, we have therefore $N = 102$ intervals over the period $T$. The series of times is: $t_0 = 8:00$, $t_1 = 8:05$, ... $t_{N−1} = 16:25$, $t_N = 16:30$.

We recall, from Chapters 2 and 4, that the element $v_k$ of the trading schedule $v$ is the volume we trade between time $t_k$ and time $t_{k+1}$. Also, $p_k$ is the log-midprice of the stock at time $t_k$.

To model the market impact of the trading list $v$ we use the propagator model in real time. We express the log-price of the stock at time $t_n$ with eq. 3.4, which we report here:

$$p_n = p_0 + \sum_{k=0}^{n-1} [\eta_k + f(v_k)G_0(n-k)].$$

(8.1)

1 The results for the modeling of a selling strategy are the same, i.e. in our model the impact costs are symmetric with respect to the trade sign.
The price of the stock is assumed to be influenced by the trading of the schedule $v$ via the terms $f(v_k)G_0(n-k)$. The trading activity of the other market participants, instead, is not included in the market impact contribution.

We are making the same assumption of [26], [27] and [28]: the impact of the activity of other traders is accounted for by the random term $\eta_k$. We now concentrate on the impact function $f$ and the propagator $G_0$.

### 8.1.1 Impact function

We need to be careful when defining the impact function $f(v)$. We recall that in 7.1.3 we studied the impact of normalized volume imbalances $v_{n_{\text{nor}}}$, obtaining:

$$f(v_{n_{\text{nor}}}) = E[r_n|v_{n_{\text{nor}}}] = \theta v_{n_{\text{nor}}}. \quad (8.2)$$

However, in order to express the impact of the volume $v_k$, we need to model the normalized volume imbalance it causes.

We first define the series $W_n$ of the total volumes traded in the market. We have that the total volume $W_k$ traded between time $t_k$ and time $t_{k+1}$ is given by:

$$W_k = \sum_{t^i \in [t_k, t_{k+1}]} |v_i^H| \quad (8.3)$$

where $v_i^H$ refers to the volumes of the single transactions. This quantity equals the denominator of the equation 7.3, which defined the normalized volume imbalances.

We can assume that the volume $v_k$ of the trading list causes a normalized imbalance of:

$$v_{k_{\text{nor}}} = \frac{v_k}{W_k}, \quad (8.4)$$

the volume traded $v_k$ divided by the total volume $W_k$. Therefore, the impact function $f$ of the volume $v_k$ is:

$$f(v_{k_{\text{nor}}}) = \frac{\theta}{W_k} v_k \equiv \theta_k v_k. \quad (8.5)$$

which defines the series $\theta_k$. It coincides with the definition we gave in equation 4.9.

In the following we will approximate the series $W_n$ with a constant value, $W_k \equiv W$. We note that, in fact, the series of total volumes traded in the market $W_k$ has a clear intraday pattern, the so-called volume profile. We will not model it because the numerical minimization would become too complicated.

The parameter $\theta$ was estimated empirically, we report its values in table 8.1.

We also define a strictly related series, the participation rate $x_k = x$, a well known concept in the trading business. Our trading activity between time $t_k$ and time $t_{k+1}$ has a participation rate of $x_k$ equal to:

$$x_k \equiv \frac{v_k}{W_k}. \quad (8.6)$$

It is the ratio between $v_k$ and the total market volume traded in the interval $W_k$. 

8.2. OPTIMAL EXECUTION WITHOUT SPREAD COSTS

8.1.2 Impact propagator

We use the functional form \ref{eq:impact_propagator} for the impact propagator $G_0$, which we report here:

$$G_{0}^{f_{t}}(t) = \frac{\Gamma_{0}}{(l_{0}^2 + t^2)^{\beta/2}}. \quad (8.7)$$

In section \ref{sec:impact_propagator_estimation} we estimated empirically the parameters of the propagator $G_0$. We show in table \ref{tab:impact_propagator_parameters} the results for the four stocks we studied.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>$\theta$ (bp)</th>
<th>$\Gamma_0$</th>
<th>$l_0$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AZN</td>
<td>15.4 ± 0.5</td>
<td>1.40 ± 0.04</td>
<td>20 ± 1</td>
<td>0.190 ± 0.006</td>
</tr>
<tr>
<td>VOD</td>
<td>26.0 ± 0.8</td>
<td>1.07 ± 0.01</td>
<td>4 ± 1</td>
<td>0.075 ± 0.002</td>
</tr>
<tr>
<td>BP</td>
<td>7.9 ± 0.3</td>
<td>1.6 ± 0.1</td>
<td>0 ± 2</td>
<td>0.21 ± 0.03</td>
</tr>
<tr>
<td>BARC</td>
<td>12.2 ± 0.5</td>
<td>1.4 ± 0.1</td>
<td>0 ± 1</td>
<td>0.18 ± 0.02</td>
</tr>
</tbody>
</table>

Table 8.1: Parameters of the fits of the market impact propagator model in the real time setting, which we use in the impact model for the optimal execution. The parameters of $G_0$ are those obtained from the fits shown in figure \ref{fig:impact_propagator_fit}.

8.2 Optimal execution without spread costs

We start with the simplest specification of the problem. Following the theory of section \ref{sec:optimal_execution}, we neglect the contribution of bid-ask spread costs.

We want to minimize the expected value $E[c(v)]$ of the log transaction costs, which are given by the expression of eq. \ref{eq:transaction_costs} reported here:

$$c(v) \equiv \sum_{k=0}^{N-1} v_k (\tilde{p}_k - p_0). \quad (8.8)$$

We recall that the effective log-midprice $\tilde{p}_k$ is the logarithm of the average mid-price at which we assume to trade the shares $v_k$ between time $t_k$ and time $t_{k+1}$, since we neglect the bid-ask spread. We note that the price $p$ of the stock is in fact a continuous time process, which we sample every 5 minutes to obtain the discrete price series $p_n$. Therefore, it is reasonable to assume that log midprice $\tilde{p}_k$ of the trade between time $t_k$ and $t_{k+1}$ is given by the average of the prices at the two times:

$$\tilde{p}_k = \frac{p_k + p_{k+1}}{2}. \quad (8.9)$$

We want to write a formula for the effective price $\tilde{p}_n$ we just obtained. Thus, we recall the effective propagator $\tilde{G}_0$ introduced in section \ref{sec:effective_propagator} which we define as:

$$\tilde{G}_0(0) = \frac{G_0(1)}{2}, \quad \tilde{G}_0(1) = \frac{G_0(1) + G_0(2)}{2}, \quad \tilde{G}_0(2) = \frac{G_0(2) + G_0(3)}{2}, \ldots \quad (8.10)$$
where the propagator $G_0$ is in the form we discussed in section \textsection{8.1.2}. Equations \textsection{8.9} \textsection{8.10} and \textsection{8.11} imply an expression for the effective price:

$$\hat{p}_n = p_0 + \sum_{k=0}^n [\eta_k + f(v_k)\tilde{G}_0(n - k)] = p_0 + \sum_{k=0}^n [\eta_k + \theta v_k \tilde{G}_0(n - k)]. \tag{8.11}$$

Therefore the expected value of the fractional transaction costs $c(v)$ is given by:

$$E[c(v)] = \sum_{n=0}^{N-1} v_n \sum_{k=0}^n \theta v_k \tilde{G}_0(n - k) = v^T \mathcal{I} v, \tag{8.12}$$

where the impact matrix $\mathcal{I}$ was defined in eq. \textsection{4.10}, i.e.

$$\mathcal{I}_{i,j} \equiv \begin{cases} \theta_i \tilde{G}_0(i - j) & i \geq j \\ 0 & i < j \end{cases} \tag{8.13}$$

Substituting $\theta_k \equiv \theta/S$ and the effective propagator $\tilde{G}_0$ of eq. \textsection{8.10}. The optimal solution $v^*$ that minimizes the expression of eq. \textsection{8.12} was found to be:

$$v^* = \frac{X}{1^T \mathcal{I}^{-1} 1} \mathcal{I}^{-1} 1, \tag{8.14}$$

which is a simple linear equation. We computed the optimal solutions $v^*$ for the four stocks we studied in the last Chapter. We plot in figure \textsection{8.1} a strictly related variable, the participation rate $x$ defined in equation \textsection{8.6}. It is proportional to the trading list, $x = \frac{v}{S}$. We recall that the trading list $v$ must satisfy:

$$\sum_{k=0}^{N-1} v_k = \frac{1}{100} \sum_{k=0}^{N-1} W_k = \frac{N \cdot W}{100}, \tag{8.15}$$

because we buy 1% of the total volume traded on the market during the day. Therefore, the participation rate $x$ satisfies:

$$\sum_{k=0}^{N-1} x_k = \frac{N}{100} \implies \langle x_k \rangle = 1\% \tag{8.16}$$

We see that the participation rate of figure \textsection{8.1} oscillates between positive and negative values, but its average value is indeed 1%.

The optimal solution consists in alternately buying and selling shares. When an investor buys he drives the price up. Thus, if he sells right away he takes advantage of the higher price and moves it back down. We plot in figure \textsection{8.2} the expected price movements due to the market impact of the optimal list $v^*$.

The optimal solution consists in weighting the alternating buys and sells so that the average participation rate is 1%. In this fashion, one completes the execution strategy paying the minimum possible impact costs\footnote{We note that the total impact costs are in any case positive, so that one does not profit from market impact, as studied in \cite{44}.}.
8.2. OPTIMAL EXECUTION WITHOUT SPREAD COSTS

Figure 8.1: Optimal execution strategies to buy 1% of daily volume, neglecting the contribution of bid-ask spread. We plot the required participation rate over each 5 minutes interval. The expected execution costs per share (of eq. 4.4) are 9.67 bp for VOD, 4.20 bp for AZN, 3.62 bp for BARC and 2.31 bp for BP.
Figure 8.2: Expected movement of the price caused by the impact of the optimal list $\mathbf{v}^*$ of section 8.2 with respect to the opening price $p_0$. 
However, the solutions of figure 8.1 are of little practical interest. In the real world, one does not trade at the mid-price of the stock, but buys at the ask price and sells at the bid (because we are considering execution with market orders). By buying, one therefore pays the mid-price plus half the bid-ask spread. By selling, one gets the mid-price minus half the bid-ask spread. Thus, alternately buying and selling bears substantial costs of bid-ask spread.

We see in the next section how including the bid-ask spread costs changes the problem.

### 8.3 Optimal execution with constant bid-ask spread

We now model the problem of optimal execution including the cost of bid-ask spread. We follow the theory developed in section 4.3. We note that we are approximating the bid-ask spread with a constant value $2 \times \delta$ throughout the day. In fact, the bid-ask spread follows a stochastic process, with strong intraday variations. Its modeling is a challenging problem we will not deal with in this work.

As we discussed at the end of last section, the contribution of bid-ask spread is crucial in determining the optimal trading lists. We recall the definition of $\delta$ in eq. 4.18, half the bid-ask spread as a fraction of the stock price. The expected value of the fractional transaction costs is therefore:

$$E[c(v)] = v^T I v + \delta 1^T |v|,$$

(8.17)

where $1$ is the vector whose elements are all 1. The impact matrix $I$ is the same as in section 8.2, the parameter $\delta$ was instead estimated empirically:

$$\delta_{AZN} = 5.27 \text{ bp}, \quad \delta_{VOD} = 10.12 \text{ bp}, \quad \delta_{BP} = 1.54 \text{ bp}, \quad \delta_{BARC} = 2.35 \text{ bp}. \quad (8.18)$$

As we noted in section 4.3, we cannot express, in a closed form, the solution $v^*$ that minimizes 4.20. The term $\propto 1^T |v|$ is not linear, therefore we can not use linear algebra to obtain the optimal solution. We instead applied a numerical optimization method, using the algorithm $L-BFGS-B$ (see [45], [46]) from the optimization routines of SciPy [40]. The resulting optimal participation rates $x^*$ are shown in figure 8.3. We recall that $x^* \propto v^*$.

We can see that including the bid-ask spread contribution in the execution costs has radically changed the form of the optimal solution obtained. The participation rate $x^*$ is positive over all the market day: if one has to pay a bid-ask spread cost on every transaction it is not convenient to sell and buy back shares.

We can see that the optimal solutions are $U$-shaped: high participation rates at market opening, low during the central hours and high at the end. This happens because the market impact propagator $G_0$ decays in time.

In fact, the high participation rate at the opening impacts much the price. Then the period of low activity lets the market absorb the impact and recover a more favorable price. The high participation rate at the end impacts the price in the future, outside our time horizon $T$. In figure 8.4 we show the expected movements of the prices $p_n$ with respect to the opening prices $p_0$, caused by the impact of the list $v^*$.

---

3 We note that we are assuming that the bid ask spread is constant during the day, while in general it is not. In fact, the bid-ask spread follows a stochastic process with a clear intraday pattern.
Figure 8.3: Optimal execution strategies to buy 1% of daily volume, considering the contribution of bid-ask spread. We plot the required participation rate over each 5 minutes interval. The expected cost of impact per share (of eq. 4.4) is 9.76 bp for VOD, 4.29 bp for AZN, 3.64 bp for BARC and 2.33 bp for BP.
8.3. OPTIMAL EXECUTION WITH CONSTANT BID-ASK SPREAD

Figure 8.4: Expected movement of the price caused by the impact of the optimal list $v^*$ of section 8.3 with respect to the opening price $p_0$. 
We compare the optimal U-shaped solutions with the one obtained by Obizhaeva and Wang [28] (shown in figure 2.3) using the exponentially decaying form of the propagator. Their solution is in a sense similar, with high trading at the very beginning and end, and low in between. However, it shows a very sharp discontinuity between the high and low trading regimes, while the solutions of figure 8.3 vary smoothly over the market day.

We also compare the optimal solution $v^*$ just obtained with the flat solution of Bertsimas and Lo [26] (section 2.3.1). It consists of constant $v_k$ over all the intervals of the period $T$. This flat solution is still standard practice in the trading industry, and therefore it is interesting to measure how much the optimal U-shaped solution outperforms it.

In table 8.2 we propose a summary of the execution costs of the optimal solutions analyzed, in particular the spread costs per share $\bar{c}_{sp}(v)$ and the impact costs per share $\bar{c}_{imp}(v)$, given by:

$$\bar{c}_{sp}(v) = \frac{\delta 1^T |v|}{|1^T v|}, \quad \bar{c}_{imp}(v) = \frac{v^T I v}{|1^T v|}.$$ (8.19)

We show these costs for all four stocks and three different solutions: the flat solution of Bertsimas and Lo [26], the U-shaped solution and the oscillating solution of section 8.2.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Execution costs per share, 1% average daily volume</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fractional impact costs (bp)</td>
</tr>
<tr>
<td></td>
<td>flat solution</td>
</tr>
<tr>
<td>AZN</td>
<td>4.36</td>
</tr>
<tr>
<td>VOD</td>
<td>9.82</td>
</tr>
<tr>
<td>BP</td>
<td>2.36</td>
</tr>
<tr>
<td>BARC</td>
<td>3.69</td>
</tr>
</tbody>
</table>

Table 8.2: We compare the fractional cost of impact and fractional cost of spread, per share, for three solutions: the flat solution of Bertsimas and Lo [26], the U-shaped solution and the oscillating solution of section 8.3 and the oscillating solution of section 8.2.

We can see that if we are concerned with the minimization of the impact costs alone, the oscillating solutions are indeed the best ones. However, when we consider the contribution of bid-ask spread, the oscillating solutions become much less attractive: by selling shares and buying them back one incurs in very high spread costs.

Instead, the flat and U-shaped solutions have fractional spread costs per share equal to $\delta$ (because all elements of $v$ have the same sign).

The comparison between the U-shaped solution and the flat one shows that the former has impact costs between 1 and 2% lower. This feature is of great appeal, both from an academic and practical point of view. As we mentioned, those flat solutions are still standard practice in the trading industry.
8.4 Efficient frontier of optimal execution

In the following we extend the model of the previous section with the effects of volatility and risk aversion. We follow the theory we developed in section 4.4, which mimicked the construction of Almgren and Chriss [27] reviewed in 2.5.

We recall in particular the coefficient of risk aversion \( \lambda \geq 0 \), and the variance matrix \( \mathcal{V} \) of equation 4.23, which gives the total variance of the execution costs. We report equation 4.24 for the execution costs with the risk aversion contribution:

\[
c_{\text{tot}}(\mathbf{v}) = \mathbf{v}^T [\mathbf{I} + \lambda \mathcal{V}] \mathbf{v} + \delta \mathbf{1}^T |\mathbf{v}|. \tag{8.20}
\]

To express the variance matrix \( \mathcal{V} \) we need to know the variance \( \sigma^2 \) of the random shocks \( \eta_n \). We compute it empirically as:

\[
\bar{\sigma}^2 = \left( \frac{\sum_{n=0}^{L} \eta_n^2}{L} \right) / L, \tag{8.21}
\]

recalling that \( L \) is the total number of intervals in the dataset. The resulting average variances \( \bar{\sigma}^2 \) for the four stocks are:

\[
\begin{align*}
\bar{\sigma}^2_{\text{AZN}} &= 350.81 \text{ (bp)}^2, \\
\bar{\sigma}^2_{\text{VOD}} &= 764.52 \text{ (bp)}^2, \\
\bar{\sigma}^2_{\text{BP}} &= 157.61 \text{ (bp)}^2, \\
\bar{\sigma}^2_{\text{BARC}} &= 326.52 \text{ (bp)}^2.
\end{align*}
\]

We now can run the same numerical optimization routine we described in the last section, minimizing the objective function 8.20. The result, for the stock BARC, is shown in figure 8.5 (and many more are in Appendix A). We plot optimal participation rates \( \mathbf{x} \) with different values of risk aversion \( \lambda \), quoting in the caption the expected costs of execution per share and their standard deviation.

As the coefficient \( \lambda \) increases, we can see that the U-shaped solutions become more and more front loaded, so that more trading activity is concentrated at the market opening rather than later in the day. This happens because, as we discussed in section 2.5.1, by trading sooner one is less exposed to the fluctuations of the price and therefore risks less. We consider the class of optimal solutions with risk-aversion proposed by Almgren and Chriss [27]. They have the functional form 2.26, and are shown in figure 2.2. We report them here:

\[
v_k^{\text{A-C}} = A \cosh(\beta(T - t_k)), \tag{8.22}
\]

where \( A \) is a normalization constant, and the parameter \( \beta \) is related to the coefficient of risk aversion \( \lambda \) by \( \beta \propto \sqrt{\lambda} \). We note that these solutions were obtained with assumptions similar to ours: constant bid ask spread costs and linear impact of the volumes traded (with constant coefficient, see eq. 2.19). Only the impact model changes. Almgren and Chriss used the model with temporary and permanent impact reviewed in section 2.4, we use the propagator model with parameters estimated empirically on the data.

\[\text{footnote}{\text{4}} \text{ We note that the variance } \sigma^2 \text{ has in fact an intraday profile: there are intervals (e.g. the ones near the closing time) whose returns have statistically higher variance than the returns over other intervals. We will not model this effect.}\]
Figure 8.5: The effect of risk aversion on the optimal trading strategy to buy 1% of daily volume, for the stock BARC (the others are similar). We show the strategy that minimizes eq. 8.20 for four different values of $\lambda$. We obtain the following results:

- Expected costs $E[c(v)] = 3.64$ bp and standard deviation $\sqrt{V[c(v)]} = 99.28$ bp for $\lambda = 0.0$,
- Expected costs $E[c(v)] = 3.79$ bp and standard deviation $\sqrt{V[c(v)]} = 71.16$ bp for $\lambda = 0.8$,
- Expected costs $E[c(v)] = 3.99$ bp and standard deviation $\sqrt{V[c(v)]} = 58.21$ bp for $\lambda = 1.6$,
- Expected costs $E[c(v)] = 4.15$ bp and standard deviation $\sqrt{V[c(v)]} = 49.65$ bp for $\lambda = 2.4$. 
8.4. EFFICIENT FRONTIER OF OPTIMAL EXECUTION

We want to compare the solutions \( v_k^{A-C} \) to the optimal solutions we obtained. In the case \( \lambda = 0 \), Almgren and Chriss’ solution is flat (it suffices to substitute \( \beta = 0 \) in eq. 8.22). We already know that the \( \lambda = 0 \) optimal solutions (of section 8.3) outperforms the flat solution by 1-2% in impact costs. But what happens with \( \lambda > 0 \)?

A handy way to compare these two classes of solutions is to draw a expected costs vs. variance plot, like the one of figure 2.1. We generate the optimal solutions of the two classes, letting \( \lambda \) vary in a given interval, compute their expected costs and variance, and draw the corresponding point in the plot. The result is shown in figure 8.6. The dots correspond to Almgren and Chriss’ solutions, the stars to the solutions resulting from our minimization procedure.

We see that our optimal solutions outperform the \( v_k^{A-C} \) solutions on the whole range of parameters of risk aversion \( \lambda \) we consider. For any given level of variance \( \sigma^2 \), our optimal solution has lower expected impact costs than the corresponding \( v_k^{A-C} \) of an amount that is roughly constant (at about 1-2%).
Figure 8.6: Optimal frontier to buy 1% of daily volume, for the four stocks we consider. Every star corresponds to one of the optimal trading schedules depicted in figures A.12 to A.36. The dots are instead the solutions proposed by Almgren and Chriss in [27], of equation 2.26. In the plots of VOD and AZN we have $\lambda \in [0, 1]$. In the plots of VOD and AZN we have instead $\lambda \in [0, 2.4]$. On the $x$ axis we have the variance $\sigma^2$ of the execution costs, and on the $y$ axis the expected value of the execution costs (per share).
Conclusion

We now highlight the most interesting results we obtained, and propose some outlooks for future work. Our original research was concentrated on two topics: the empirical estimation of the market impact and the definition of optimal trading strategies. In both cases we worked within the market impact propagator model.

The estimation procedure for the propagator $G_0$ we have developed in Chapter 5 has proven very effective. When we applied it in Chapter 7 we obtained compelling results, with all three different models for time aggregation. By comparing the $R^2$ coefficients of the regressions we gained interesting insights into how the market mechanisms have evolved over the last years, a feature that can be investigated much further. In particular, one could run the estimation method on a greater number of stocks, from different markets, to see how the different models perform.

When estimating $G_0$ we have used trade time, aggregated trade time and real time, but there are many other possible choices for time aggregation. For example one could choose volume aggregation: time is increased by one unit each time a certain aggregate volume is traded on the market. It is a well known definition in the trading industry.

In addition, the empirical calibration of optimal executions of Chapter 8 in the theoretical framework of Chapter 4 has lead to very appealing results. We have seen that the optimal strategies obtained from the minimization procedure lead to cost savings of the order of 1-2% with respect to the industry standard. Also, when considering risk-aversion, we have the same margin of cost savings over a large portion of the efficient frontier of optimal solutions.

There are, however, many possible extensions one could investigate. For example we have assumed that the impact function is constant over the day, neglecting the well-observed intraday pattern of volume traded. That inclusion would most likely change the form of the solutions and put much more computational stress on the numerical minimization routine.

The intraday patterns of bid-ask spread and volatility have been neglected as well. In fact, as we argued in the text, those are stochastic processes and their modeling is a challenging problem of its own. By including their contribution, however, one could build a much richer model. Lots of effort by researchers in the industry is indeed directed to the modeling of bid-ask spread dynamics.

Throughout this work we assumed that the execution is performed only by market orders, but in fact one could post limit orders as well. A common sense practice would be to use preferably limit orders when the spread is wide and market orders otherwise. The problem of splitting between limit and market orders is therefore entrenched with the modeling of bid-ask spread dynamics.
We also mention that in the recent years the trading industry has witnessed the birth of many new trading venues, mainly electronic exchanges. Some of those venues operate in a completely different way than the limit order book markets we described. Dark pools, for instance, are potentially very interesting for the problem of optimal execution.

In this Thesis we have seen how a phenomenological model introduced in the econophysics literature, the market impact propagator, can serve as a flexible and powerful tool to determine the optimal execution strategies, a problem of great practical interest. Even more interestingly, we have estimated the costs of the optimal strategies obtained with this “physicist” approach, and found that they are better than the standard practice in the trading industry.
Appendix A

Optimal schedules with risk aversion

A.1 2000-2002 dataset, flat volume profile

![Diagram](image)

Figure A.1: Optimal schedule with risk aversion $\lambda = 0.0$. The expected cost is 9.76 bp with standard deviation 151.10 bp for VOD, and 4.29 bp with standard deviation 102.01 bp for AZN.
Figure A.2: Optimal schedule with risk aversion $\lambda = 0.1$. The expected cost is 9.77 bp with standard deviation 137.26 bp for VOD, and 4.28 bp with standard deviation 95.04 bp for AZN.

Figure A.3: Optimal schedule with risk aversion $\lambda = 0.2$. The expected cost is 9.82 bp with standard deviation 121.70 bp for VOD, and 4.30 bp with standard deviation 93.40 bp for AZN.
**A.1. 2000-2002 DATASET, FLAT VOLUME PROFILE**

![Graph](image)

Figure A.4: Optimal schedule with risk aversion $\lambda = 0.3$. The expected cost is 9.88 bp with standard deviation 112.16 bp for VOD, and 4.31 bp with standard deviation 86.02 bp for AZN.

![Graph](image)

Figure A.5: Optimal schedule with risk aversion $\lambda = 0.4$. The expected cost is 9.94 bp with standard deviation 104.43 bp for VOD, and 4.33 bp with standard deviation 82.25 bp for AZN.
Figure A.6: Optimal schedule with risk aversion $\lambda = 0.5$. The expected cost is 10.00 bp with standard deviation 98.05 bp for VOD, and 4.36 bp with standard deviation 78.88 bp for AZN.

Figure A.7: Optimal schedule with risk aversion $\lambda = 0.6$. The expected cost is 10.05 bp with standard deviation 92.69 bp for VOD, and 4.38 bp with standard deviation 75.86 bp for AZN.
A.1. 2000-2002 DATASET, FLAT VOLUME PROFILE

Figure A.8: Optimal schedule with risk aversion $\lambda = 0.7$. The expected cost is 10.10 bp with standard deviation 87.06 bp for VOD, and 4.41 bp with standard deviation 73.13 bp for AZN.

Figure A.9: Optimal schedule with risk aversion $\lambda = 0.8$. The expected cost is 10.15 bp with standard deviation 82.92 bp for VOD, and 4.44 bp with standard deviation 70.65 bp for AZN.
Figure A.10: Optimal schedule with risk aversion $\lambda = 0.9$. The expected cost is 10.20 bp with standard deviation 79.29 bp for VOD, and 4.47 bp with standard deviation 68.40 bp for AZN.

Figure A.11: Optimal schedule with risk aversion $\lambda = 1.0$. The expected cost is 10.25 bp with standard deviation 76.09 bp for VOD, and 4.50 bp with standard deviation 66.33 bp for AZN.
A.2 2011 January-April Dataset, Flat Volume Profile

Figure A.12: Optimal schedule with risk aversion $\lambda = 0.0$. The expected cost is 3.64 bp with standard deviation 99.28 bp for BARC, and 2.33 bp with standard deviation 67.50 bp for BP.
Figure A.13: Optimal schedule with risk aversion $\lambda = 0.1$. The expected cost is 3.66 bp with standard deviation 92.56 bp for BARC, and 2.33 bp with standard deviation 64.53 bp for BP.

Figure A.14: Optimal schedule with risk aversion $\lambda = 0.2$. The expected cost is 3.65 bp with standard deviation 91.02 bp for BARC, and 2.34 bp with standard deviation 62.33 bp for BP.
Figure A.15: Optimal schedule with risk aversion $\lambda = 0.3$. The expected cost is 3.67 bp with standard deviation 87.09 bp for BARC, and 2.35 bp with standard deviation 60.69 bp for BP.

Figure A.16: Optimal schedule with risk aversion $\lambda = 0.4$. The expected cost is 3.69 bp with standard deviation 83.69 bp for BARC, and 2.34 bp with standard deviation 60.76 bp for BP.
Figure A.17: Optimal schedule with risk aversion $\lambda = 0.5$. The expected cost is 3.71 bp with standard deviation 80.68 bp for BARC, and 2.35 bp with standard deviation 59.20 bp for BP.

Figure A.18: Optimal schedule with risk aversion $\lambda = 0.6$. The expected cost is 3.74 bp with standard deviation 75.98 bp for BARC, and 2.36 bp with standard deviation 57.78 bp for BP.
Figure A.19: Optimal schedule with risk aversion $\lambda = 0.7$. The expected cost is 3.76 bp with standard deviation 73.46 bp for BARC, and 2.37 bp with standard deviation 56.46 bp for BP.

Figure A.20: Optimal schedule with risk aversion $\lambda = 0.8$. The expected cost is 3.79 bp with standard deviation 71.16 bp for BARC, and 2.38 bp with standard deviation 55.24 bp for BP.
Figure A.21: Optimal schedule with risk aversion $\lambda = 0.9$. The expected cost is 3.81 bp with standard deviation 69.06 bp for BARC, and 2.39 bp with standard deviation 54.09 bp for BP.

Figure A.22: Optimal schedule with risk aversion $\lambda = 1.0$. The expected cost is 3.84 bp with standard deviation 67.13 bp for BARC, and 2.40 bp with standard deviation 53.02 bp for BP.
A.2. 2011 JANUARY-APRIL DATASET, FLAT VOLUME PROFILE

Figure A.23: Optimal schedule with risk aversion $\lambda = 1.1$. The expected cost is 3.86 bp with standard deviation 65.35 bp for BARC, and 2.41 bp with standard deviation 52.02 bp for BP.

Figure A.24: Optimal schedule with risk aversion $\lambda = 1.2$. The expected cost is 3.89 bp with standard deviation 63.71 bp for BARC, and 2.42 bp with standard deviation 51.07 bp for BP.
Figure A.25: Optimal schedule with risk aversion $\lambda = 1.3$. The expected cost is 3.91 bp with standard deviation 62.18 bp for BARC, and 2.43 bp with standard deviation 50.18 bp for BP.

Figure A.26: Optimal schedule with risk aversion $\lambda = 1.4$. The expected cost is 3.94 bp with standard deviation 60.77 bp for BARC, and 2.44 bp with standard deviation 49.33 bp for BP.
A.2. 2011 JANUARY-APRIL DATASET, FLAT VOLUME PROFILE

Figure A.27: Optimal schedule with risk aversion $\lambda = 1.5$. The expected cost is 3.96 bp with standard deviation 59.44 bp for BARC, and 2.45 bp with standard deviation 48.54 bp for BP.

Figure A.28: Optimal schedule with risk aversion $\lambda = 1.6$. The expected cost is 3.99 bp with standard deviation 58.21 bp for BARC, and 2.46 bp with standard deviation 47.78 bp for BP.
Figure A.29: Optimal schedule with risk aversion $\lambda = 1.7$. The expected cost is 4.01 bp with standard deviation 57.05 bp for BARC, and 2.48 bp with standard deviation 45.49 bp for BP.

Figure A.30: Optimal schedule with risk aversion $\lambda = 1.8$. The expected cost is 4.02 bp with standard deviation 55.28 bp for BARC, and 2.49 bp with standard deviation 44.76 bp for BP.
A.2. 2011 JANUARY-APRIL DATASET, FLAT VOLUME PROFILE

Figure A.31: Optimal schedule with risk aversion $\lambda = 1.9$. The expected cost is 4.05 bp with standard deviation 54.21 bp for BARC, and 2.50 bp with standard deviation 44.08 bp for BP.

Figure A.32: Optimal schedule with risk aversion $\lambda = 2.0$. The expected cost is 4.07 bp with standard deviation 53.19 bp for BARC, and 2.52 bp with standard deviation 43.42 bp for BP.
Figure A.33: Optimal schedule with risk aversion $\lambda = 2.1$. The expected cost is 4.09 bp with standard deviation 52.24 bp for BARC, and 2.53 bp with standard deviation 42.79 bp for BP.

Figure A.34: Optimal schedule with risk aversion $\lambda = 2.2$. The expected cost is 4.11 bp with standard deviation 51.33 bp for BARC, and 2.54 bp with standard deviation 42.19 bp for BP.
Figure A.35: Optimal schedule with risk aversion $\lambda = 2.3$. The expected cost is 4.13 bp with standard deviation 50.47 bp for BARC, and 2.55 bp with standard deviation 41.61 bp for BP.

Figure A.36: Optimal schedule with risk aversion $\lambda = 2.4$. The expected cost is 4.15 bp with standard deviation 49.65 bp for BARC, and 2.56 bp with standard deviation 41.06 bp for BP.
Bibliography


