The Thermodynamics of High Frequency Markets

Kevin Thomas Webster

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Operations Research and Financial Engineering
Adviser: Professor Carmona

June 2014
Abstract

High Frequency Trading (HFT) represents an ever growing proportion of all financial transactions as most markets have now switched to electronic order book systems. This dissertation proposes a novel methodology to analyze idiosyncrasies of the high frequency market microstructure and embed them in classical continuous time models.

The main technical result is the derivation of continuous time equations which generalize the self-financing relationships of frictionless markets to electronic markets with limit order books. We use NASDAQ ITCH data to identify significant empirical features such as price impact and recovery, rough paths of inventories and vanishing bid-ask spreads. Starting from these features, we identify microscopic identities holding on the trade clock, and through a diffusion limit argument, derive continuous time equations which provide a macroscopic description of properties of the order book.

These equations naturally differentiate between trading via limit and market orders. We give several applications to illustrate their impact and how they can be used to the benefit of Low Frequency Traders (LFTs). In particular, option pricing and market making models are proposed and solved, leading to new insights as to the impact of limit orders and market orders on trading strategies.
Acknowledgements

First and foremost, I would like to thank my adviser, René Carmona, for accepting the research problem I gave myself nearly four years ago and helping me stick through with it throughout the PhD. The angles of attack and tools we chose changed from year to year, but his insightful advice and passion stayed constant and strong.

Research is never done completely in isolation, and I am grateful to Professors Yacine Ait-Sahalia, Jean Jacod and Ramon van Handel for advice on limit theorems for path functionals, a tool that ended up being crucial in my dissertation. Professors Markus Brunnermeier and Patrick Cheridito, both helped nudge my research in the right direction during and after my generals examination. They also imparted important knowledge on modeling heterogeneous information and convex analysis respectively. I am also happy to have shared my office with Geoffrey Zhu, Michael Li, Haifeng Luo and especially Daniel Lacker. We spent a lot of time together on control problems, mean field games, data analysis and just plain beautiful Mathematics.

Other people whom I have interacted with and who provided some form of inspiration for this research are Professors Alexander Schied, Robert Almgren, Jakob Jurek, Ronnie Sircar, Birgit Rudloff, Sasha Stoikov along with seminar participants of the Oxford-Princeton, Cambridge-Princeton and Humboldt-Princeton workshops. The ORFE department and the Bendheim institute in general have been a fruitful and welcoming research environment with plenty of helpful interactions.

Finally, the original impulse for my research stemmed from an internship at BNP Paribas doing high frequency market making in Foreign Exchange. I would like to thank all the members of the team, and in particular Valentin Clément, Tibor Gergely, Sebouh Takvorian and Renaud Drappier for sparking my interest.
To my parents
Contents

Abstract ................................................................. iii
Acknowledgements ...................................................... iv

0 Introduction .......................................................... 1

1 Review ................................................................. 6
   1.1 The limit order book ............................................ 6
   1.2 Macroscopic models ............................................ 7
      1.2.1 The self-financing equation ............................ 8
      1.2.2 Our basic formula ........................................ 9
      1.2.3 Almgren and Chriss .................................... 10
      1.2.4 Transaction cost literature ............................ 11
   1.3 Micro-to-macro models of the limit order book ............ 12
      1.3.1 Cont and de Larrard ................................... 13
      1.3.2 Horst and Paulsen ..................................... 14

2 Microscopic theory .................................................. 15
   2.1 Formalizing the limit order book ............................ 15
      2.1.1 Basic relationships .................................... 16
      2.1.2 Alternative representation of the order book ....... 20
      2.1.3 Wealth .................................................. 24
      2.1.4 Summary ................................................ 25
## 5 Applications

5.1 Option pricing .................................................. 82
   5.1.1 Mathematical setup ........................................ 82
   5.1.2 The result .................................................. 84

5.2 Absence of price manipulation .................................. 86

5.3 Market making .................................................. 87
   5.3.1 Microscopic model .......................................... 88
   5.3.2 Control problem on the macroscopic limit .............. 89
   5.3.3 Solution .................................................... 90

## 6 A macroscopic model for market making in foreign exchange markets

6.1 Context and self financing equation .......................... 97
   6.1.1 Market making in Foreign Exchange .................... 97
   6.1.2 A different diffusion limit ............................... 98

6.2 Client control problem ......................................... 100
   6.2.1 Heterogeneous beliefs ..................................... 100
   6.2.2 Trade assumptions ......................................... 102
   6.2.3 The control problem ....................................... 103
   6.2.4 Dynamics of the implied alpha .......................... 106
   6.2.5 Robustness analysis ....................................... 108

6.3 Defining the market maker’s control problem ............... 109
   6.3.1 Approximate price process ................................. 110
   6.3.2 Approximate objective function .......................... 114
   6.3.3 Infinitely many clients .................................... 115
   6.3.4 Modeling the error term ................................... 119

6.4 The market maker’s control problem ........................ 120
   6.4.1 Model summary ............................................ 120
Chapter 0

Introduction

In a series of papers ([23] [24] [25]) on the divide between high and low frequency traders, M. O’Hara and co-authors identified a number of market features that both Low Frequency Traders (LFTs for short) and most academic researchers have largely ignored, but that High Frequency Traders (HFTs from now on) exploit with great success.

“There is no question that the goal of many HFT strategies is to profit from LFTs mistakes. [...] Part of HFTs success is due to the reluctance of LFT to adopt (or even to recognize) their paradigm.” ([25])

These papers also outline a program to better understand and possibly remedy these issues: in a nutshell, these authors recommend that LFTs update the strategies and models they use in order to incorporate more of the features of the high frequency markets. While the goal should not be to try to beat the HFTs at their own game by modeling the high frequency market microstructure in painstaking detail, it should be to capture, at least sparsely, the macroscopic effects of those phenomena that actually affect LFT.

The dissertation is in line with this program. Case in point, its main thrust is to provide forms of the self-financing portfolio equation, both in discrete and continuous
time, consistent with the high frequency paradigm. The equations we propose are motivated by and fitted to high frequency data. They are derived theoretically from accounting rules at the high frequency level. Their continuous time limits capture the relevant effects at the macroscopic level. From these fundamental relationships, we use the powerful tools of stochastic calculus to revisit the solutions of a certain number of standard continuous time financial problems in light of the new high frequency paradigm. We show how the latter affects for example option hedging and we highlight the differences depending upon trading being through limit orders versus market orders. A model for market making in the spirit of [9] is solved. We also prove results regarding the absence of price manipulation strategies under certain microstructure assumptions.

The crucial insight of [25], named 'the new paradigm', is the fact that high frequency traders do not operate on the 'calendar' clock, but instead use some form of 'event-based time', such as the trade clock, or the volume clock. This is partly due to the algorithmic nature of their strategies and the lack of direct calendar clock dependent constraints such as maturities and the likes. A fringe benefit for quantitative analysis is the well documented fact that prices behave better under an event-based clock than the calendar clock. A number of papers [8, 16, 17, 36, 37, 25] argue that, in addition to removing seasonal effects and resolving asynchronicity issues, this time-change makes the price returns more Gaussian-like. Even though this property is mostly irrelevant in our analysis, we choose to work in the trade clock in which each discrete time step corresponds to one trade. Indeed, even though our conclusions are independent of the clock used, we find the trade clock especially convenient to formulate and test the significance of our findings.
With these proviso out of the way, we can outline our research agenda:

1. Understand, at the microscopic level, structural relationships and strategies that HFTs exploit;

2. Identify which features persist at the macroscopic level, in which form, and provide continuous time models on that scale;

3. Use these models to update LFT strategies.

For the sake of definiteness, we focus on the self-financing portfolio equation of continuous time finance. To this effect, we review in chapter 1 the role of this condition in quantitative finance, and in so doing, introduce the continuous time analysis notation used in the paper, as well as the exact form of our generalization.

The main originality of the form of the self-financing condition which we propose to use, is the fact that it accounts for both price impact and price recovery, two important empirical microstructure features that are usually ignored or modeled in separate ad-hoc fashions. It also differentiates between the impacts of limit and market orders. This is important because nowadays, a large number of agents trade with both types of orders, rather than simply relying on market makers to find trades. Furthermore, our generalization of the self-financing portfolio equation can be used with a larger class of inventories models, e.g. with infinite variation. This allows the use of the powerful tools of stochastic calculus to retain tractability in a number of models.

The classical self-financing portfolio equation was generalized in two separate directions in the financial engineering literature. On one hand, Almgren and Chriss proposed in [6] a way to incorporate price impact and temporary transaction costs in a phenomenological model for optimal execution with market orders and finite speed of trading. On the other hand, and with a completely different point of view, extensions of the classical self-financing equation of the Black-Scholes theory were touted
by researchers attempting to include transaction costs in Merton’s optimal portfolio’s theory. See for example [15, 35, 45] or the recent review [34]. Both start off in the macroscopic limit and we hope that our approach of modeling the microscopic behavior of markets first will gain traction in the mathematical finance community. Note that queuing-theory based models, such as [18] already exhibit this behavior and apply it at the order book level, but not on actual trading strategies.

Two books, Empirical market microstructure by J. Hasbrouck ([27]) and Market microstructure theory by M. O’Hara ([42]) cover the state of the field prior to the advent of HFT. They contain informed trader models ([33]) and inventory-based market making ([7, 26, 28, 43]). Three main themes united different market structures at that time: the limit order book, adverse selection (the underlying cause of price impact) and statistical predictions. These themes are just as relevant, if not more so in the new age of high frequency trading.

Our investigations were inspired by a large number of empirical studies of high frequency data (see for example [10, 11, 12, 14, 39, 48, 49, 50]), and recent publications of theoretical models of the limit order book ([18, 19, 20, 29, 38, 49]). However, our emphasis is different as we use limit orders as a starting point. Our goal is not to explain the evolution of the order book, but merely to analyze the consequences of the choices made by the liquidity providers and takers on price changes, their inventories and their wealth.

The rest of the dissertation is structured in the following way:

Chapter 1 reviews basic notions of high frequency microstructure before going over the current state of the literature on the topic. In particular, self-financing equations used in mathematical finance are analyzed and compared to the one we propose. The chapter ends with diffusion limits of limit order book models, which follow a similar logic to our proposed methodology for the self-financing equation.
Chapters 2 to 4 are the crux of the thesis, as they implement our research agenda. Chapters 5 and 6 illustrate the results on applications and hence provide the closure to our program.

Chapter 2 is our microscopic description of the market. It begins with a formalization of the limit order book. This then allows the introduction of equations and models for high frequency microstructure phenomena described earlier.

We then study real high frequency data in the light of our microscopic models in chapter 3. The objective is twofold: use the models to provide monitoring tools for the market, as well as justifying and testing said models and crucial assumptions used later on.

The main result of the thesis is presented in chapter 4: a new self-financing portfolio equation consistent with high frequency data. This equation includes transaction costs and price impact, also called adverse selection. It allows for inventories of infinite variations and distinguishes between trading with limit or market orders. All these features allow it to perfectly track real life wealth while remaining tractable and opening up models related to limit order or market order strategies.

Chapter 5 presents applications of this self-financing portfolio equation. The Black and Scholes formula is updated in a setting with transaction costs and price impact and we derive a result on delta-hedging with limit orders. Then price manipulation is explored in the case of a flat order book. Finally, a first model of market making is presented.

The thesis ends with a more elaborate market making model in chapter 6. This was the original motivation of the author, and particular care is given to adapt it to a particular market: Foreign Exchange. This chapter also serves to showcase a derivation and application of a different, more standard self-financing equation.
Chapter 1

Review

In this section we review some facts and models about high frequency markets. Section 1.1 describes trading on a limit order book. Previous self-financing portfolio equations are covered in section 1.2. We also introduce our proposed self-financing equation for comparison. Finally, we mention the literature revolving around diffusion limits for queuing models of the limit order book in section 1.3 as these present some common points with our proposed methodology.

1.1 The limit order book

Trading on high frequency markets takes place on an object called the limit order book. An agent can interact with others via two possible trading mechanisms: limit orders and market orders. Limit orders correspond to the act of providing liquidity to the market, while market orders take liquidity from it. We will refer to agents who engage in the first type of trade as liquidity providers\(^1\) while traders who trade with market orders will be referred to as liquidity takers. In real markets, traders often switch between liquidity providing and taking strategies, blurring this definition somewhat. The following comments can help highlight the differences.

\(^1\)Of which market makers are a special class.
• A liquidity taker pays a fee for his aggressiveness. This fee typically takes the form of the bid-spread, which is where most trades happen. The corresponding provider captures this bid-ask spread.

• Right after the trade happens, the price may move. If it does, it almost always moves in favor of the market order, compensating to some degree the transaction costs. This phenomena is called price impact. It is a consequence of the adverse selection of limit orders by takers.

• Between two successive trades, the price reverts to some value in between the impacted price and the original one. Price recovery is an intuitive name often used to describe this high frequency feature.

• Takers control their inventory directly. Attaining correlation with the market requires high frequency predictions of the next price move.

• Providers do not directly control their inventory, but only their exposure to the flow of market orders. How much of the flow they are able to capture depends on their limit order fill rate. Flow is considered toxic if it leads to adverse selection. The profitability of a provider’s strategy depends on the spread he captures and the toxicity of his flow.

1.2 Macroscopic models

The current approach in mathematical finance is to model the market in continuous time. The main reason to do so is the tractability of models involving stochastic calculus. An empirical justification for using Itô processes to model prices and other financial quantities is the sheer number of trades happening on markets, making financial graphs look -at least on the hourly scale- like rough but continuous paths: exactly like those of a regular stochastic process.
We adhere to this motto, albeit with a twist: our emphasis is to justify continuous
time equations and models by building up more fundamental equations and models
on the microscopic, trade-by-trade scale before obtaining their continuous time coun-
terpart by a limiting procedure. The main reason to do this is that, unfortunately,
different sets of assumptions lead to different macroscopic models. We now review
the standard continuous time models with respect to our own to accentuate different
features of macroscopic modeling.

1.2.1 The self-financing equation

In quantitative finance, the standard self financing portfolio equation is a cornerstone
of the theory of frictionless markets. It plays a crucial role in many fundamental re-
sults, e.g. Merton’s portfolio theory. Mathematically speaking, it is a simple equation
which constrains the wealth process of an investor to live in a certain sub-space. This
sub-space is therefore often called the space of admissible portfolios. New-comers
to the mathematical theories of financial market often gripe with the self-financing
condition and how it relates to the real world. While it can be postulated as a
mathematical definition, it can also be derived from a limiting procedure starting
from accurate descriptions of the microstructure of trades in the trade clock. This
approach is at the core of our strategy.

“The sad fact is that the self-financing condition is considerably more
subtle in continuous time than it is in discrete time.” ([16])

When discussing market models at the macroscopic level, we assume that the
mid-price \( p \) and the inventory \( L \) are given by Itô processes:

\[
\begin{align*}
    dp_t &= \mu_t dt + \sigma_t dW_t \\
    dL_t &= b_t dt + l_t dW'_t
\end{align*}
\]

(1.1)
for two Wiener processes $W$ and $W'$ with unspecified correlation structure. In the simplest case, we also consider an adapted process $s_t$ representing (in the continuous time limit) the bid-ask spread measured \textit{in tick size}.

The standard self-financing condition of continuous time finance can be stated as a constraint:

$$dX_t = L_t dp_t$$

between the price $p$ of the underlying interest, the inventory $L$, and the wealth $X$ of the agent. In most classical financial applications, e.g. Merton’s portfolio theory, the price $p$ is exogenously given, the inventory $L$ is the agent’s input, and his wealth $X$ appears as the output of equation (1.2).

The objective of this paper is to generalize the self-financing portfolio condition (1.2) to incorporate known idiosyncrasies of the high frequency markets including transaction costs, price impact and price recovery. Also, we want this generalization to be able to quantify the differences between trading via limit orders and market orders. We warn the reader that the equations proposed in this dissertation are only necessary conditions and that quantifying limit order fill rates, priorities and price recovery are beyond the immediate scope of the present paper.

### 1.2.2 Our basic formula

In the case where trades happen only at the best bid or ask, we formulate the self-financing condition in the following form:

$$dX_t = L_t dp_t \pm s_t t \frac{1}{\sqrt{2\pi}} dt + d[L,p]_t$$

where $\pm$ is $+$ when trading with limit orders, and $-$ when trading with market orders. Indeed, we show in chapter 4 that, when time is measured in the trade clock, the discrete time analog of formula (1.3) can be derived rigorously from a specific
limit order book feature. It also matches real wealth data as we will see in chapter 3. We shall also impose the constraint

\[ d[L, p] < 0 \] (1.4)

whenever trading with limit orders. The interpretation for this constraint is price impact and has also been thoroughly tested on the data in chapter 3.

We now explain how our condition (1.3) and the adverse selection constraint (1.4) relate to the conditions used in the literature.

1.2.3 Almgren and Chriss

The seminal work by Almgren and Chriss [6] addresses a question closely related to ours. These authors propose a macroscopic model for the price impact and the change of wealth after a liquidity taker’s decision. The model leads to a very tractable framework which was used by many optimal execution studies (see [3, 41] for example). This framework can be summarized by the system:

\[
\begin{align*}
    dp_t &= f_t(l_t)dt + \sigma_t dW_t \\
    dL_t &= l_t dt \\
    dX_t &= L_t dp_t - c_t(l_t)dt
\end{align*}
\] (1.5)

where \( f \) and \( c \) are two function-valued adapted processes which are positive, and in the case of \( c \), convex.

The main advantage of this model is that price impact appears in a tractable fashion. Indeed, it comes through the function \( f_t \), which creates a positive ‘correlation’ between traded volumes and the price process. However, it constrains \( L \) to be differentiable and for this reason, the model parameters cannot be calibrated to
market data directly, making the model difficult to test empirically. As our empirical analysis in chapter 3 there is ample evidence supporting inventories with infinite variations (see also [18]). Moreover, limit orders are not part of the discussion in the Almgren-Chriss framework. We recover the self financing equation of Almgren and Chriss in chapter 6 for a particular scaling of the order book.

1.2.4 Transaction cost literature

The branch of classical mathematical finance most related to our paper is portfolio selection under transaction costs ([15, 35, 45] or the recent review [34]). Most of these works start from a model for the wealth of a liquidity taker which generalizes the self-financing equation to a setting with transaction costs. In general however, these papers do not emphasize the derivation of the model, but instead, the study of its consequences. We hope to appeal to this side of the community by providing more accurate equations for self-financing portfolios while keeping similar tractability, leading the way to problems related to liquidity provision, such as market making. An interesting feature of such problems is that the agent does not directly control his portfolio, adding an additional modeling challenge. For the record we note that the standard equation used in this branch of the literature is

\[ dX_t = L_t dp_t - \frac{s_t}{2} |dL_t| \]

where again, the inventory process \( L \) is assumed to have finite variation \( \int_0^t |dL| < \infty \) for all finite \( t \) and \( s_t \) is the bid-ask spread.

Strengths of this model are its simplicity, relative tractability, and straightforward calibration to the market. Its weaknesses include the fact that the process \( L \) can only have finite variation. Moreover, price impact, limit orders and other microstructure considerations are absent in the model.
Formula (1.6) is much closer to our proposed equation (1.3) than it may seem at first. It merely corresponds to a different diffusion limit. It can be recovered in our framework by considering non-vanishing bid-ask spread, zero price impact and looking at market orders only. Notice that these assumptions may be more natural than ours for low frequency markets. This is presumably the reason for their introduction.

1.3 Micro-to-macro models of the limit order book

Our main objective is to derive the self-financing equation from a microscopic description of the market. This relies heavily on market microstructure and a keen understanding of the limit order book. The dividend is the resolution of certain continuous financial applications in the presence of such microstructure frictions. We believe the incorporation of such market frictions is worth the added effort to go from a microscopic to a macroscopic model rather than directly postulating a macroscopic model.

Other authors have pursued such a micro-to-macro approach for the modeling of the limit order book. The starting point is a discrete time queuing model of the limit order book. A scaling argument is then used to obtain a macroscopic limit via either a functional law of large numbers or a functional central limit theorem. Our methodology coincides with this branch of the financial mathematics literature, although our objective differs. Our goal is to study trading strategies and we therefore focus on the self-financing portfolio equation, while he goal of the following two papers is to study the statistical properties of the limit order book.
1.3.1 Cont and de Larrard

The article [18] by Cont and de Larrard builds on the seminal paper [20] by the former. [20] is the first modern paper applying queuing theory to study statistical properties of the limit order book. To gain tractability, the model assumes Poisson arrival rates. [18] circumvents this hypothesis by considering a diffusion approximation. This uses the fact that vastly different time scales are present on high frequency markets.

A reduced form microscopic model

The paper starts with a description of the dynamics of the order book on the microscopic scale. This allows for an exact tracking of the limit orders arrivals, cancellations and executions. The proposed model is of reduced form type, which allows for both tractability and easier empirical analysis, as all the modeled quantities are observable on markets.

The scaling argument

A crucial question Cont and de Larrard ask is which limiting regime to use, as this will dictate their choice of scaling parameters for the limit order book. Using high frequency data, they observe that order flow exhibits diffuse behavior on the macroscopic scale and the authors use this insight to justify their diffusion limit.

Macroscopic limit

A functional central limit theorem then leads to the final macroscopic model for the limit order book and price dynamics. The final macroscopic model can be summarized by bid and ask queue lengths that follow a reflected 2-dimensional Wiener process with reflections inside the positive quadrant.
Common points

Our methodology goes through the same three steps: a microscopic model that exactly tracks all the relevant quantities, a data analysis to justify the scaling regime for our diffusion limit and finally the application of a functional limit theorem to derive the macroscopic equations. The diffusive behavior noted in [18] is also present in our data for inventories, leading to what we believe is a very similar macroscopic description of the market, even though we focus our modeling effort on different objects than Cont and de Larrard.

1.3.2 Horst and Paulsen

The paper by Horst and Paulsen [29] follows a similar idea: modeling the limit order book on the microscopic scale before going through with a limiting argument. The main differences with [18] are:

- A more elaborate microscopic model, which is more structural in nature and describes the full limit order book rather than focusing on the best bid or ask levels.
- A different scaling regime which leads to ODE and PDE systems based on a functional law of large numbers.
Chapter 2

Microscopic theory

Borrowing from a time-honored method in statistical physics, we first describe in depth the interactions between agents at a microscopic level. This will pave the way for our data analysis as well as the macroscopic diffusion limit argument.

This chapter begins in section 2.1 with a mathematical description of the limit order book and related trade rules. This description takes as given the limit orders present prior to a trade and the incoming trading volume and defines or derives the market quantities resulting from the trade. This includes the cash exchanged, the price impact of the trade and the change in wealth of both parties of the trade. Extensions to multiple agents and the special case where all trades happen at the best bid or ask price are explored in section 2.2. A microscopic description of price impact and adverse selection follows in section 2.3, first using a reduced form approach before proposing a structural model.

2.1 Formalizing the limit order book

We consider a single liquidity taker and a single liquidity provider. They trade an asset whose possible price range is $(0, \infty)$ via a limit order system. The liquidity provider always moves first by choosing the limit orders she places on the limit order
book. These limit orders are represented by a control variable \((b,a)\) consisting of a pair of strictly positive measures on \((0,\infty)\). The liquidity taker then chooses the control variable \((\beta,\alpha) \in (0,\infty) \times (0,\infty)\) representing market orders that he wants to execute on that limit order book.

Throughout this section we use the liquidity provider’s point of view to track changes in portfolio positions and ignore the following high frequency phenomena:

1. *Slippage.* Market orders execute immediately at their intended price.

2. *Partial fills.* Market orders consume all the volume present at a given price\(^1\).

3. *Hidden orders.* All limit orders are public.

Note that we will drop assumption 2 in section 2.2.2, where we study a market in which all the trades happen either at the best bid or the best ask price.

### 2.1.1 Basic relationships

We first focus on basic relationships between the two agents, their orders and inventories.

The control \((b,a)\) of the liquidity provider represents her limit orders. A bid for one unit of the asset placed at a price \(p\) is represented by the probability measure \(b = \delta_p\), while an offer (or ask) for one unit at price \(p'\) by \(a = \delta_{p'}\). If the provider places multiple limit orders, we sum these unit masses and obtain two non-negative measures \(b\) and \(a\) representing the liquidity provider’s aggregate orders.

We will call \((b,a)\) a limit order book, or order book. We define the best bid and ask of an order book in the following way:

\[ a = 1.3 \delta_{p_1} + 0.8 \delta_{p_2} \]

\(^1\)This property automatically holds when you formally consider a continuous order book distribution.
Definition 2.1.1 (Best bid and ask). Let \((b,a)\) be an order book. Then we define the best bid and best ask prices to be

\[
\bar{b} = \sup\{p \in \text{supp}(b)\}, \quad a = \inf\{p \in \text{supp}(a)\}
\] (2.1)

Here we use the notation \(\text{supp}(\mu)\) for the topological support of the measure \(\mu\).

Remark 2.1.2. In real markets, such limit orders can only be placed on a discrete grid, and the resulting \(a\) and \(b\) are always discrete measures. The recent push of high frequency markets to refine their grid may justify considering measures \(a\) and \(b\) that are absolutely continuous with respect to the Lebesgue measure.

The control \((\beta, \alpha)\) of the liquidity taker represents his market orders. A market order placed against the bids will cause all the bid orders above and including the price \(\beta\) to be executed. For market orders against the ask, all the limit orders below the level \(\alpha\) will be executed. The limiting cases \(\alpha = 0\) and \(\beta = \infty\) correspond to 'empty' market orders that do not execute any limit orders. The execution of a market order leads to the following changes in cash and inventory:

Definition 2.1.3 (Execution of a market order). Assume the order book is \((b,a)\) and that the liquidity taker chooses the pair \((\beta, \alpha)\). Then the change \(\Delta L\) of inventory triggered by the trade and the change \(\Delta K\) in cash that the liquidity provider is subject to are defined by:

\[
\Delta L = b[\beta, \infty) - a(0, \alpha] \tag{2.2}
\]

\[
\Delta K = \int_{(0,\alpha]} xa(dx) - \int_{[\beta,\infty)} xb(dx) \tag{2.3}
\]
For the justification of this formula let us first consider a single bid \( b = \delta_p \). That is, the provider expresses interest in buying one unit of the asset at the price \( p \) or lower. A liquidity taker’s market order to sell will therefore execute the order if and only if its price level \( \beta \) is smaller. Should such an execution take place, the liquidity provider gains one unit of volume and loses \( p \) units of cash.

The above formula is then obtained by aggregating linearly the individual limit orders.

The following assumptions will be used.

**Assumption 2.1.4.** The order books \( (b, a) \) are such that \( \bar{b} < a \), that is, the bid-ask spread is always positive. We will say in this case that the order book exhibits no arbitrage.

**Assumption 2.1.5.** It is never optimal for the liquidity taker to buy and sell simultaneously.

In particular, we can recode the liquidity taker’s control by a single real number \( \alpha \) by making him formally send the market orders \((\alpha, \alpha)\). Indeed, if \( \alpha \in (\bar{b}, a) \) there is no trade, if \( \alpha \geq a \) a buy happens but no sell, and conversely for \( \alpha \leq \bar{b} \).

**A probabilistic model for liquidity taker behavior**

We now provide a simple model for which Assumption 2.1.5 follows from Assumption 2.1.4. Let \( (\Omega, \mathcal{F}, \mathbb{P}) \) be a probability space modeling the beliefs of the liquidity taker. Let \( p \) be a random variable representing the price at which the liquidity taker values the asset at a future time. We assume that the liquidity taker is risk-neutral under \( \mathbb{P} \) in the sense that he maximizes his expected wealth after the trade. In other words
he solves the optimization problem:

$$\max_{\beta, \alpha} \mathbb{E}[-p\Delta L - \Delta K]$$

(2.4)

**Proposition 2.1.6.** Let the order book \((a, b)\) be given. Then an optimal trade for the liquidity taker is given by

$$\beta = \alpha = \mathbb{E}[p].$$

(2.5)

**Proof.** The liquidity taker looks for the supremum over \((0, \infty) \times (0, \infty)\) of the function

$$(\beta, \alpha) \mapsto \int_{[\beta, \infty)} (x - \mathbb{E}[p])b(dx) - \int_{(0, \alpha]} (x - \mathbb{E}[p])a(dx).$$

(2.6)

This function decouples and we are left maximizing

$$\beta \mapsto \int_{[\beta, \infty)} (x - \mathbb{E}[p])b(dx)$$

(2.7)

which is non-decreasing on \((0, \mathbb{E}[p])\) and non-increasing on \([\mathbb{E}[p], \infty)\). The same result holds for

$$\alpha \mapsto -\int_{(0, \alpha]} (x - \mathbb{E}[p])a(dx).$$

(2.8)

The supremum is attained for \(\beta = \alpha = \mathbb{E}[p]\). \qed

**Remark 2.1.7.** While we do not have uniqueness of this maximum, all the other choices of optimum market orders will lead to exactly the same executions. Indeed, the function \(\beta \mapsto \int_{[\beta, \infty)} (x - \mathbb{E}[p])b(dx)\) and \(\alpha \mapsto -\int_{(0, \alpha]} (x - \mathbb{E}[p])a(dx)\) respectively do not have a strict maximum in \(\mathbb{E}[p]\) iff \(b\) and \(a\) respectively put zero mass on some interval including \(\mathbb{E}[p]\). Any market orders on this interval will lead to exactly the same cash and asset transfers and we can without loss of generality replace them by market orders at \(\mathbb{E}[p]\). In particular, we can summarize the taker’s market orders by a single number \(\alpha\).
Corollary 2.1.8. Assume the order book \((b,a)\) exhibits no arbitrage. Then it is never optimal for the taker to buy and sell simultaneously.

Proof. By the previous comment, we can summarize the market orders of a taker behaving optimally by a single real \(\alpha\). The taker’s buy and sell volumes are

\[
a[a, \alpha] \text{ and } b[\alpha, \bar{b}] \tag{2.9}
\]

The no arbitrage property implies that these two terms cannot both be positive. \(\square\)

2.1.2 Alternative representation of the order book

Even though the above representation of limit and market orders is clear, we still present an alternative description which only makes sense if no arbitrage is present on the market and Assumption 2.1.5 is verified.

The below definitions correspond to a very intuitive ‘graphic’ approach. In section 2.1.1 we have defined the order book as a pair of positive measures \((b,a)\). The no-arbitrage condition guarantees that these two measures have disjoint supports. One is therefore tempted to ‘glue’ the two measures together into one. But in order to do that, we also need to keep track of where the offers starts and the bids stop. This is done in the following way.

Definition 2.1.9 (Quoted price). Let \((b,a)\) be an order book that does not exhibit arbitrage. Then we say that \(p\) is a quoted price of the order book if \(p \in (\bar{b}, \bar{a})\).

Because the bid-ask spread is positive, there is not a unique quoted price. This is an unfortunate reality of high frequency markets, and we will only be able to mathematically resolve this difficulty in the limit where the bid-ask spread vanishes. Using a quoted price as a separation point between bid and ask limit orders, we can define:
Definition 2.1.10 (Shape function). Let \((b,a)\) be an order book that exhibits no arbitrage and \(p\) be one of its quoted prices. Then define the order book’s shape function \(\gamma : \mathbb{R} \mapsto [0, \infty)\) to be

\[
\gamma(u) = \int_0^u (a(0,p+x) - b(p+x,\infty)) \, dx.
\]  

(2.10)

In particular, \(\gamma\) is convex, \(\gamma(0) = 0\) and \(\gamma'(0) = 0\). Moreover, if the measures \(a\) and \(b\) have finite mass, \(\gamma'\) is bounded and as a result, \(\gamma\) has at most linear growth.

Remark 2.1.11. Notice that \(\gamma''(\cdot+p) = b + a\) if both measures \(b\) and \(a\) have densities, or more generally, if we understand this equality in the sense of distributions.

The gamma function

---

21
The following result recasts the trade equations in terms of the function $\gamma$.

**Proposition 2.1.12.** Let $(b, a)$ be an order book which exhibits no arbitrage, $p$ be one of its quoted prices and $\gamma$ the associated shape function. If $\alpha = u + p$ is the liquidity taker’s market order, then we have

$$\Delta L = -\gamma'(u)$$
(2.11)

$$\Delta K = (u + p)\gamma'(u) - \gamma(u).$$
(2.12)

**Proof.** The first identity is immediate from the definition of $\gamma$ and $\Delta L$:

$$\Delta L = b[\alpha, \infty) - a(0, \alpha]$$

$$= -\gamma'(u)$$

The second identity follows using integration by parts:

$$\Delta K = \int_{(0,\alpha]} xa(dx) - \int_{[\alpha,\infty)} xb(dx)$$

$$= \alpha a(0, \alpha] - \int_{(0,\alpha]} a(0, x)dx - \alpha b[\alpha, \infty) + \int_{[\alpha,\infty)} b[x, \infty)dx$$

$$= \alpha \gamma'(u) - \gamma(u)$$

\[\square\]

**Remark 2.1.13.** Our assumption on fill rates is what allows us to perfectly reconstruct the volume triggered by a market order at price level $\alpha$. In the general case, if there is a point mass on the order book at price level $\alpha$ and market orders are not required to take all of the volume, any volume $\Delta L \in (\gamma'(u+), \gamma'(u-)]$ is permissible.
To relate the transaction costs back to the traded volume without going through the transaction price $\alpha$, we use the following result:

**Proposition 2.1.14. (Transaction costs)** Define the transaction cost function $c$ as the Legendre transform of $\gamma$:

$$c(l) = \sup_u (ul - \gamma(u)). \tag{2.13}$$

Then we have:

$$\Delta K = -p\Delta L + c(-\Delta L), \tag{2.14}$$

and in particular, $c$ is convex and satisfies $c(0) = 0$.

**Proof.** By the Fenchel identity, we have that

$$u\gamma'(u) = \gamma(u) + c(\gamma'(u))$$

and that $c'$ is the generalized inverse of $\gamma'$. Hence, as $\Delta L = -\gamma'(u)$ we have that $u = c'(-\Delta L)$ and

$$\Delta K = -p\Delta L + u\gamma'(u) - \gamma(u)$$

$$= -p\Delta L + c(-\Delta L)$$

\[ \square \]

An order book $(b, a)$ which does not exhibit arbitrage can therefore be represented by a pair $(p, \gamma)$ with $p$ a real and $\gamma$ a differentiable, convex function satisfying $\gamma(0) = \gamma'(0) = 0$. If the total mass of the limit order book is finite, the function $\gamma$ has linear growth. Note that this representation in terms of quoted price and order book shape is not unique, but leads to a completely equivalent description of trades and hence the same market model.
Both representations have pros and cons and unfortunately, both will need to be juggled at different times of our analysis. The advantages of the original \((b,a)\) representation are: uniqueness of the decomposition, ease to derive no-arbitrage relationships and natural interpretation of formulas. The alternative representation in terms of \((p,\gamma)\) is more tractable and concise as it involves a real number and a function rather than a pair of measures.

2.1.3 Wealth

Wealth is a quantity of interest to traders summarizing a cash and asset position in a single number. This is done by fixing a quoted price \(p\), usually the midpoint between the bid and the ask price, and using cash as a numeraire. This leads to the definition

**Definition 2.1.15.** Let \(p\) be a quoted price, \(L\) a liquidity provider’s inventory and \(K\) her cash account. We define by \(X_p\) the wealth marked to the quoted price \(p\) with the formula

\[
X_p = pL + K
\]

(2.15)

In the case where \(p = (\bar{b} + a)/2\) is the midpoint between the bid and the ask, we drop the index \(p\) and call it wealth marked to the midprice.

We can use the formulas derived before to write out the change in wealth given a choice of quoted prices \(p\) and \(p + \Delta p\) for both time periods. This change in wealth can be written as:

\[
\Delta X = (p + \Delta p)(L + \Delta L) - pL + \Delta K
\]

\[
= L\Delta p + p\Delta L + \Delta p\Delta L + \Delta K
\]

\[
= L\Delta p + c(-\Delta L) + \Delta p\Delta L
\]

where the first line is a definition resulting from (2.15).
This is our first self-financing portfolio equation. A general feature of these equations when trading frictions such as transaction costs are included is the fact that wealth becomes a non-linear functional of the trading path. This is what will motivate our use of functional law of large numbers and central limit theorems in chapter

2.1.4 Summary

For future reference, we summarize the different trade equations and market representations defined and derived. These equations happen to also hold in special cases we derive later on, although care must be taken when rederiving them.

The liquidity provider places limit orders. If the limit order book formed that way presents no arbitrage, it will be represented either by a pair of measures \((b, a)\) or a couple \((p, \gamma)\) with \(p\) a real number and \(\gamma\) a differentiable, convex function with linear growth and \(\gamma(0) = \gamma'(0) = 0\). Consistency equations between the two representations can be found above.

We call \((b, a)\) the order book, \(p\) a quoted price and \(\gamma\) the shape of the order book. The liquidity taker’s market order will be represented either by a real \(\alpha\) representing a price, or a real \(u\) denoting a centered price (shifted by the quoted price \(p\)). Both representations lead to the same trades.

\[
\Delta L = b[\alpha, \infty) - a(0, \alpha] = -\gamma'(u)
\]
is the change in inventory of the liquidity provider, while
\[
\Delta K = \int_{(0, \alpha]} x a(dx) - \int_{[\alpha, \infty)} x b(dx)
\]
\[
= (u + p)\gamma'(u) - \gamma(u)
\]
\[
= p\Delta L + c(-\Delta L)
\]
is her change in cash position.

The market order corresponding to this trade can be recovered from the limit orders and the trade volume by the relationship

\[
\alpha - p = c'(-\Delta L) \tag{2.16}
\]

and the change in wealth if the quoted price moves by \(\Delta p\) is

\[
\Delta X = L\Delta p + c(-\Delta L) + \Delta p\Delta L \tag{2.17}
\]

### 2.2 Additional cases

The equations summarized in section 2.1.4 were derived in the case where a single liquidity provider trades with a single liquidity taker and trades consume all the limit orders at the traded price. While these equations also hold in more general cases, care must be taken when deriving or using the formulas.

#### 2.2.1 The case of multiple providers

In this subsection, we will cover the case where multiple agents are present on the market. We forgo modeling any form of priority or speed advantage by assuming that takers never trade simultaneously, which makes it completely equivalent to there only
being one liquidity taker. As for providers, we wish to make use of the 'aggregate limit order book', which requires us to abstract away the queuing priority of limit orders. This is automatically the case under our assumptions, as trades always trade all the volume available at a given price level.

We consider a finite number of liquidity provider indexed by \( j = 1 \ldots n \) for some \( n > 0 \).

**Definition 2.2.1.** Let \( (b_j, a_j)_{j=1}^{n} \) be the limit orders associated to each liquidity provider. Then we define the aggregate limit order book to be the order book

\[
(b, a) = \left( \sum_{j=1}^{n} b_j, \sum_{j=1}^{n} a_j \right)
\]  

(2.18)

This definition leads to the following aggregate shape and transaction cost functions.

**Proposition 2.2.2.** Let \( p \) be a quoted price of the aggregate order book \( (b, a) \) associated to the individual order books \( (b_j, a_j)_{j=1}^{n} \). Then \( p \) is a quoted price of each individual order book and if \( (p, \gamma_j) \) is the associated dual representation, then

\[
\gamma = \sum_{j=1}^{n} \gamma_j
\]  

(2.19)

that is, the aggregate order book shape is simply the sum of the individual shapes with respect to a common quoted price. Furthermore, the aggregate transaction cost function is obtained by inf-convolution of the individual transaction cost functions.

\[
c = \Box_{j=1}^{n} c_j
\]  

(2.20)

**Proof.** Because the measures \( b_j \) and \( a_j \) are all non-negative, the support of each \( b_j \) is included in the support of \( b \) and similarly for the asks. This leads to \( \bar{b} \geq \bar{b}_j \) and \( \underline{a} \leq \underline{a}_j \) for each \( j \). In words, the aggregate bid-ask spread is always at least as narrow
as each individually quoted bid-ask spread. Any quoted price of the aggregate limit
order book is therefore within each individual bid-ask spread.

The formula for the aggregate shape function follows from the linearity of the
integration operation, while the formula for the aggregate transaction cost function
follows from the Legendre duality between the sum and inf-convolution operations.

The trading equations follow:

**Corollary 2.2.3.** For a liquidity taker trading volume $\Delta L$, the formulas from section
[2.1.4] hold by using the aggregate limit order book, shape function and transaction cost
function.

For a liquidity provider quoting the orders $(b_j, a_j)$, we have that

$$\Delta L_j = b_j[\alpha, \infty) - a_j(0, \alpha]$$

$$= -\gamma_j'(u)$$

and

$$\Delta K_j = \int_{(0, \alpha]} xa_j(dx) - \int_{[\alpha, \infty)} xb_j(dx)$$

$$= (u + p)\gamma_j'(u) - \gamma_j(u)$$

$$= p\Delta L_j + c_j (-\Delta L_j)$$

where the market order $\alpha$ can be recovered from the total trade volume via the formula

$$\alpha - p = c'(-\Delta L) \quad (2.21)$$
2.2.2 Special case: the bid-ask spread

In this section we assume that trades happen only at the best bid or ask but drop the assumption that trades always consume all the liquidity available at the traded price. The latter means that the traded volume is not only a function of the traded price anymore. If a trader uses market orders, it is easy to quantify which trade volumes are admissible: he can trade any quantity smaller or equal to the volume available on the best bid (respectively ask) when selling (respectively buying). For limit orders, this depends on the size of the incoming trade on the priority system used by the exchange. The most common priority system is a First In, First Out (FIFO) queue.

In terms of modeling, it means we cannot assume as given the market order price $\alpha$ and back out from it the traded volume. Instead, our primary data is directly the traded volume $\Delta L$. Once this quantity is given, the previous formulas summarized in 2.1.4 hold. This in particular means that we can use as a self-financing equation

$$\Delta X = L\Delta p + \frac{s}{2}|\Delta L| + \Delta p\Delta L$$

(2.22)

where $s$ is the bid-ask spread, $p$ is the midprice and we assume the trader is using limit orders.

2.3 Price impact and adverse selection

The self-financing equation derived in 2.1.4 can be considered as a bare bones relationship and corresponds to an accountant’s perspective on the market. Given a trader’s inventory and the limit order book he or she trades on, the accountant can track his or her wealth perfectly. The number of hypotheses made are minimal to obtain the result and it perfectly tracks wealth once a trading strategy is given.
The above framework does not pronounce itself on which trading strategies are admissible however. Clearly, if any strategy is permissible, trading with limit orders is always preferable to trading with market orders as you capture the transaction costs rather than pay them. But in practice there is a trade-off when choosing to trade via limit or market orders. This trade-off has been captured by key words such as adverse selection, price impact and market response function. These three terms all are related on an informal level and correspond to attempts by different communities (the Economics, Math Finance and Econophysics community respectively) to model this hidden cost of placing limit orders. We will propose our own approach within the framework built up in the previous two sections.

2.3.1 Price impact and recovery

The terminology 'price impact' stems from the move of the midprice right after a trade. Indeed, if the midprice before a trade is \( p \), the trade is of size \( \Delta L \) from the limit order’s perspective and the transaction costs are a continuous function \( c \), then, right after the trade, the midprice will move from \( p \) to \( p + c'(-\Delta L)/2 \) as either the bid or the ask will move by \( c'(-\Delta L) \).

However, in between trade times a lot of actions happen on the limit order book, some of which may change the price. The change in price after the price moves from a trade is called 'price recovery'. It is stochastic and relatively hard to predict. On average, it acts as a dampener on the price impact, making it revert to some point in between the pre-trade price and the impacted price.

While one can compute explicitly the impacted price right after the trade, price recovery must be modeled as it is not a mechanical relationship. A model for the overall effect of a trade on the midprice must therefore be proposed. The seminal paper by Almgren and Chriss (6) is an important example of such a model. We propose two alternatives.
2.3.2 A reduced form adverse selection constraint

In words, our reduced form price impact model claims that the price always moves in favor of market orders. Within our framework, this corresponds to the following hypothesis.

**Assumption 2.3.1.** Let \( p \) be the midprice, \( L \) the inventory of a liquidity provider. Then we have that

\[
\Delta p \Delta L \leq 0 \quad (2.23)
\]

**Interpretation**

While our model is purely descriptive and agnostic to agent-based models, we can still interpret the adverse selection within the framework of Kyle’s model [33]. Consider a high frequency exchange with two types of agents placing market orders: noise traders and informed traders.

In Kyle’s model an informed trader is a trader who knows the value of the asset at the end of the trading period and trades accordingly. His inventory will be correlated with the price at the end of the day, but the changes in his inventory are potentially uncorrelated with the changes in prices on a trade-by-trade basis.

Because we model a high frequency exchange, our notion of informed trader differs from Kyle’s. In such a setting, the traders most dangerous to a market maker or any agent placing limit orders are not the ‘regular’ traders with an accurate medium or long-term view on the market. The traders liquidity providers are worried about are algorithmic traders who predict with near perfect accuracy the next price movement. They tend to not have a longer term view on the market and attempt to make a round-trip trade within a very short time window. These traders therefore hold an inventory that is uncorrelated with the price at the end of the day, but whose changes are very much correlated with the changes in price on a trade by trade basis.
Figure 2.1: Illustration of different types of informed traders and how inventories can be correlated to the price in different ways: on a more long term or on a trade-by-trade basis.
In such a high frequency market, informed traders will always trigger trades such that $\Delta p \Delta L > 0$ from their perspective, that is, $\Delta p \Delta L < 0$ from the liquidity provider’s perspective. By predicting prices on a very short time horizon, these high frequency traders guarantee to never engage in a trade that will lose them money in the immediate future. Whether it will actually be profitable enough to compensate them for the transaction costs paid is a different question.

2.3.3 A structural price impact model

The reduced form model is useful because it assumes very little. This makes it robust and easily testable against the data. It does not provide a predictive model for how prices change after trades however, and one may want to impose such a structure onto the model. This makes the model move closer to the notion of ‘price impact’ as developed by Almgren and Chriss ([6]).

The proposed structural model is deterministic price recovery of the price after each trade.

**Assumption 2.3.2.** Let $\lambda \in (0, 1]$ be a real that encapsulates price recovery. The bigger $\lambda$, the smaller the price recovery. Then assume that

$$\Delta p = \lambda c'(-\Delta L)$$  \hspace{1cm} (2.24)

or, equivalently

$$\Delta L = -\gamma'(\lambda^{-1} \Delta p)$$  \hspace{1cm} (2.25)

Because we know exactly by how much the price moves immediately after a trade, imposing a deterministic recovery of the price from its impacted level creates a structural system where the price and a provider’s inventory are mechanically linked.
Chapter 3

Data analysis

The aim of this chapter is to illustrate the equations of chapter 2 on empirical high frequency data.

In section 3.1 we present our data source, the Nasdaq ITCH files. Our choices in terms of processing and trade reconstruction are detailed in this section.

Section 3.2 discusses the empirical wealth process in relationship to self-financing wealth equations. We show how previous equations lack crucial terms, and how our equation, which tracks the wealth without such errors, decomposes the trading strategy into three components with clear interpretations.

Finally, section 3.3 covers price impact and adverse selection in the data. This is the most important section of the chapter, as it 1) guides our limiting procedure by ruling out certain scaling models for the order book 2) serves as a rigorous statistical test of our reduced form model.

3.1 Presentation of the data

Our raw data consists of NASDAQ’s ITCH data files, which include all visible limit and market orders, as well as market orders that execute hidden liquidity. We process the data by removing ‘special deals’ trade that happen within the bid-ask spread and
market orders that execute hidden liquidity. This leaves us with only visible market
orders hitting visible limit orders at the best bid or ask price.

3.1.1 Processing the data

The ITCH data format is very rich and contains a lot of information absent from
other sources, such as TAQ from the Wharton Research Data Services. However,
with increased granularity also come some technical challenges. We present here our
proposed solutions and conventions regarding the matter.

Discarded trades

For the sake of completeness, we include a table of how many executions were dis-
carded for each stock. They correspond to symbols 'C' and 'P' in the NASDAQ ITCH
messaging convention. Type 'C' messages correspond to special deals while type 'P'
messages to trades that execute on a hidden limit order.

<table>
<thead>
<tr>
<th>stock</th>
<th>type 'C'</th>
<th>type 'P'</th>
<th>stock</th>
<th>type 'C'</th>
<th>type 'P'</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>1.4%</td>
<td>16.6%</td>
<td>FCN</td>
<td>0.3%</td>
<td>24%</td>
</tr>
<tr>
<td>AAPL</td>
<td>0.4%</td>
<td>36.3%</td>
<td>FFIC</td>
<td>0%</td>
<td>24.2%</td>
</tr>
<tr>
<td>ADBE</td>
<td>1.3%</td>
<td>7.5%</td>
<td>FL</td>
<td>1.4%</td>
<td>6.2%</td>
</tr>
<tr>
<td>AGN</td>
<td>1.2%</td>
<td>12.9%</td>
<td>FMER</td>
<td>1.5%</td>
<td>8.4%</td>
</tr>
<tr>
<td>AINV</td>
<td>1.4%</td>
<td>18.8%</td>
<td>FPO</td>
<td>0.2%</td>
<td>24.1%</td>
</tr>
<tr>
<td>AMAT</td>
<td>1.8%</td>
<td>14.2%</td>
<td>FRED</td>
<td>0.3%</td>
<td>7.4%</td>
</tr>
<tr>
<td>AMED</td>
<td>0.7%</td>
<td>13.6%</td>
<td>FULT</td>
<td>1.6%</td>
<td>8.5%</td>
</tr>
<tr>
<td>AMGN</td>
<td>1.5%</td>
<td>11.8%</td>
<td>GAS</td>
<td>1.1%</td>
<td>18.3%</td>
</tr>
<tr>
<td>AMZN</td>
<td>0.5%</td>
<td>25.9%</td>
<td>GE</td>
<td>2.7%</td>
<td>9.1%</td>
</tr>
<tr>
<td>ANGO</td>
<td>0.3%</td>
<td>19%</td>
<td>GILD</td>
<td>1.4%</td>
<td>9.5%</td>
</tr>
<tr>
<td>APOG</td>
<td>0.5%</td>
<td>15.2%</td>
<td>GLW</td>
<td>1.6%</td>
<td>13.4%</td>
</tr>
<tr>
<td>ARCC</td>
<td>4.5%</td>
<td>9.9%</td>
<td>GOOG</td>
<td>0.9%</td>
<td>31%</td>
</tr>
</tbody>
</table>

35
Proportion of ignored trades. Type 'C' correspond to 'special deals' while type 'P' execute against hidden limit orders.

Reconstructing parent orders

When we theoretically discuss the case where all trades happen at the best bid or ask price, we make a seemingly strong assumption on the data. Indeed, in general trades may go deeper into the book and consume liquidity from the next bid and ask prices. Unfortunately, the raw ITCH data does not indicate how to reconstruct these 'parent orders' from the single orders seen in the file. We however propose a reconstruction heuristic and present our empirical results in this general case. We also present the same data analysis without employing this heuristic. Note that in the latter case, all trades by definition of the ITCH data format happen at the best bid or ask price, as
parent orders are automatically split into child orders that must happen at the best
bid or ask price.

The heuristic we employ to reconstruct parent orders in the ITCH data is as follows. We declare that multiple trades are children of the same parent order if the
following three conditions are satisfied

1. They all have the exact same timestamp.

2. They are all in the same direction (Buy or Sell).

3. In between each trade message, there are no other messages on the same stock.
   In particular, the order book did not change in between each trade message.

We believe this to be a rather conservative heuristic, but it already yields a compres-
sion of, on average, four child trades into one parent order. Most trades still happen
at the best bid or ask price.

3.1.2 The aggregate liquidity provider

Because the data does not give us the identity of the different traders, we work with
a representative liquidity provider. This means we look at all the trades triggered
by limit orders and construct the ensuing portfolio and consider that as our 'trader'.
The figures given for illustrative purposes were produced using the data for Coca Cola
(KO) on 18/04/13. A full cross-section of 120 stocks used in the recent ECB study
was considered over multiple days to test for robustness of our results. Note that
some of the stocks were removed from certain tables when the number of trades was
not sufficiently large to run the corresponding analysis.

We first define the quantities of interest. They are all backward-looking.

Let $p_n$ be the midprice before the trade happens at time $n$. Similarly, let $s_n$ be
the bid-ask spread before the trade. Denote by $L_n$ the aggregate liquidity provider’s
inventory, that is, his net position in the traded asset before the trade at time $n$. 38
Figure 3.1: Example: Coca Cola (KO) stock on 18/04/13. Inventory, cash and wealth are those of the aggregate liquidity provider.
Finally, define by $K_n$ the amount of cash the trader is holding and assume that his position is self-financing. This hypothesis allows us to reconstruct the liquidity provider’s wealth using the relationship

$$X_n = p_n L_n + K_n$$ (3.1)

### 3.2 Wealth decomposition

Recall the discrete self-financing equation (2.22) for limit orders trading at the best bid or ask derived in section 2.2.2

$$\Delta_n X = L_n \Delta_n p + \frac{s_n}{2} |\Delta_n L| + \Delta_n L \Delta_n p$$ (3.2)

or

$$\Delta_n X = L_n \Delta_n p + c_n (-\Delta_n L) + \Delta_n L \Delta_n p$$ (3.3)

in the general case where we reconstruct parent trades.

#### 3.2.1 Discussion

Equation (3.2) is a simple accounting rule for updating the wealth of a portfolio when trading on a high frequency order book. It contains three terms, two of which have a clear interpretation:

**Frictionless wealth**

The standard equation used in a frictionless market is

$$\Delta_n X = L_n \Delta_n p$$ (3.4)
which coincides with our equation \((3.2)\) when the trader holds his position without trading. The rational for calling this the *frictionless* wealth equation is the notion that whether the trader is holding the position or re-balancing it should not affect his wealth. Whether this holds true in practice depends upon the values of the two other terms of equation \((3.2)\).

**Transaction costs**

The term

\[
\pm \frac{s_n}{2} |\Delta_n L|
\]

has a clear interpretation as the spread captured or paid during the transaction. If one trades with limit orders, the spread is captured for each unit of traded asset while for market orders, the spread is paid. This term, all others being equal, favors trading with limit orders and constitutes a first friction on high frequency markets. Note that the exact form of this transaction cost term hinges on the assumption that all trades are made at the best bid or ask price. In the general case, transaction costs are a convex function of the trade volume.

**Adverse selection**

The third term is non-standard, and to the best of our knowledge, has never been identified and studied in this direct form:

\[
\Delta_n \Delta_n L.
\]

We now propose an interpretation and a model for this term linked to adverse selection.

It is positive whenever the midprice moves in the same direction as the traded volume after the trade happens, and negative otherwise. Note the slight temporal
gap between the two terms of the product: $\Delta_n p = p_{n+1} - p_n$ is the difference between the midprice right before the next trade and right before the current trade. It is not equal to the difference between the midprice after and before the trade, as limit orders may be canceled or placed between two trade times.

However, when the midprice does not move between the trade at time $n$ and the trade at time $n+1$, we have that

$$\Delta_n p \Delta_n L \leq 0 \quad (3.7)$$

if the trade happened with limit orders and

$$\Delta_n p \Delta_n L \geq 0 \quad (3.8)$$

otherwise. This is a purely mechanical relationship of the order book. For instance, immediately after a buy market order hits a sell limit order, the ask is either moved up or stays at the same level. Never does it move down immediately due to a trade. This is called 'price impact of trades' by practitioners. It is interesting to notice that these inequalities nearly always hold, even when limit orders are being canceled or replaced between trade times. This is due to the adverse selection of limit orders by agents using market orders.

Finally, we therefore proposed in section 2.3 the model

$$\pm \Delta_n p \Delta_n L \leq 0 \quad (3.9)$$

where "±" is "+" for limit orders and "−" for market orders. In particular, this term is always of the opposite sign to the transaction cost term. Limit orders capture the spread but get adversely selected.

This model will be studied more in depth in section 3.3.
3.2.2 Empirical wealth

We now compare several formulas and models for tracking a trader’s wealth. We use the aggregate liquidity provider’s true wealth as the benchmark our wealth equation is supposed to track.

The standard way to model wealth in the academic finance literature is to use the frictionless self-financing wealth equation:

$$\Delta_n X = L_n \Delta_n p$$ (3.10)

where $p$ is not the midprice, but the ‘fair’ price of the asset, which in an efficient market, is assumed to be a martingale. This equation does not match the data when we use the midprice and the concept of 'microstructure noise' can be invoked to explain this significant gap. In a frictionless market, microstructure does not matter and profits come solely from longer term views on the market.

Transaction costs can be added, leading to the self-financing equation with transaction costs:

$$\Delta_n X = L_n \Delta_n p \pm \frac{s_n}{2} |\Delta_n L|$$ (3.11)

While this equation takes into account the spread, it ignores price impact. Is is however a good model for agents who are using market orders and who do not adversely select the market. This can be the case for low frequency traders who do not optimize their execution and only use long-term views to trade.

Finally, incorporating the price impact term leads to a complete picture of wealth on a high frequency market:

$$\Delta_n X = L_n \Delta_n p \pm \frac{s_n}{2} |\Delta_n L| + \Delta_n L \Delta_n p$$ (3.12)
This description forgoes the notion of microstructure noise, as wealth can be perfectly tracked using directly measurable market quantities.

Three idealized types of traders are often used to describe a high frequency market and correspond to strategies that involve profiting on each of the three components:

- Low frequency traders, who profit from the frictionless part of the equation by trading on longer term views.
- Market makers, who capture the transaction costs component by providing liquidity.
- Liquidity arbitrageurs, who exploit short-term signals and adversely select the market.

We first compare the impact of transaction costs and price impact by computing wealth with each of the three equations: the frictionless case, the case with transaction costs and the exact wealth equation. We provide a table with the results as well as a plot in figure 3.2 for our example stock, Coca Cola on 04/18/13. The plot in particular shows how current models of wealth cannot even track the realized wealth \textit{ex-post}! While microstructure noise models such as \cite{2} can be used to measure the error introduced by ignoring the high frequency microstructure, an exact reconstruction of wealth that matches the data perfectly is possible and not more difficult.

As expected, the frictionless equation underestimates true wealth while the equation ignoring price impact over-estimates the wealth of our aggregate liquidity provider. The relative error is always significant. This comes as no surprise as microstructure noise adds up very quickly over a trading day.

The frictionless wealth plays a different role than the two other components, which are purely microstructure related. We therefore directly compare the ratio between the third and the second component.
Figure 3.2: Example: Coca Cola on 04/18/13. Plots of the actual wealth of the aggregate liquidity provider together with the wealth computed from the three self-financing conditions. Red is the frictionless case (3.10). Green corresponds to (3.11). The actual wealth and the wealth computed from our self-financing condition (3.12) are indistinguishable on the graph.
The ratio between the money lost to instantaneous adverse selection and transaction cost suggests that adverse selection must always be taken into account. While it seems like transaction costs more than compensate the liquidity provider for instantaneous adverse selection - the ratio is always smaller than 1 - the average liquidity provider’s profits are still negative, suggesting that adverse selection also acts on longer time horizons than the next tick. Unconditional market making therefore loses money on average, but while instantaneous adverse selection is a significant part of those losses, it does not explain all of it, and some long-term adverse selection must also be factored in.

<table>
<thead>
<tr>
<th>stock symbol</th>
<th>relative error</th>
<th>friction ratio</th>
<th>net P&amp;L</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>177%</td>
<td>8%</td>
<td>5970 $</td>
</tr>
<tr>
<td>AAPL</td>
<td>9%</td>
<td>19%</td>
<td>-1779103 $</td>
</tr>
<tr>
<td>ADBE</td>
<td>49%</td>
<td>44%</td>
<td>-5746 $</td>
</tr>
<tr>
<td>AGN</td>
<td>564%</td>
<td>78%</td>
<td>-28 $</td>
</tr>
<tr>
<td>AINV</td>
<td>140%</td>
<td>16%</td>
<td>1062 $</td>
</tr>
<tr>
<td>AMAT</td>
<td>80%</td>
<td>7%</td>
<td>-15806 $</td>
</tr>
<tr>
<td>AMED</td>
<td>173%</td>
<td>40%</td>
<td>181 $</td>
</tr>
<tr>
<td>AMGN</td>
<td>15%</td>
<td>46%</td>
<td>-55395 $</td>
</tr>
<tr>
<td>AMZN</td>
<td>63%</td>
<td>30%</td>
<td>-57491 $</td>
</tr>
<tr>
<td>ANGO</td>
<td>149%</td>
<td>44%</td>
<td>127 $</td>
</tr>
<tr>
<td>APOG</td>
<td>93%</td>
<td>45%</td>
<td>947 $</td>
</tr>
<tr>
<td>ARCC</td>
<td>54%</td>
<td>21%</td>
<td>2690 $</td>
</tr>
<tr>
<td>AXP</td>
<td>19%</td>
<td>41%</td>
<td>-37162 $</td>
</tr>
<tr>
<td>AYI</td>
<td>266%</td>
<td>64%</td>
<td>495 $</td>
</tr>
<tr>
<td>AZZ</td>
<td>162%</td>
<td>50%</td>
<td>57 $</td>
</tr>
<tr>
<td>BAS</td>
<td>1%</td>
<td>96%</td>
<td>-1594 $</td>
</tr>
<tr>
<td>BHI</td>
<td>6%</td>
<td>61%</td>
<td>44602 $</td>
</tr>
<tr>
<td>Symbol</td>
<td>Expected</td>
<td>Actual</td>
<td>Difference</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>BIIB</td>
<td>51%</td>
<td>40%</td>
<td>-19266 $</td>
</tr>
<tr>
<td>BRCM</td>
<td>41%</td>
<td>29%</td>
<td>23826 $</td>
</tr>
<tr>
<td>BRE</td>
<td>8%</td>
<td>58%</td>
<td>-3772 $</td>
</tr>
<tr>
<td>BXS</td>
<td>108%</td>
<td>69%</td>
<td>-133 $</td>
</tr>
<tr>
<td>BZ</td>
<td>18%</td>
<td>85%</td>
<td>-208 $</td>
</tr>
<tr>
<td>CB</td>
<td>58%</td>
<td>58%</td>
<td>-1417 $</td>
</tr>
<tr>
<td>CBEY</td>
<td>81%</td>
<td>41%</td>
<td>399 $</td>
</tr>
<tr>
<td>CBT</td>
<td>25%</td>
<td>74%</td>
<td>769 $</td>
</tr>
<tr>
<td>CBZ</td>
<td>1235%</td>
<td>67%</td>
<td>-21 $</td>
</tr>
<tr>
<td>CDR</td>
<td>13%</td>
<td>62%</td>
<td>-260 $</td>
</tr>
<tr>
<td>CELG</td>
<td>55%</td>
<td>48%</td>
<td>-12403 $</td>
</tr>
<tr>
<td>CETV</td>
<td>122%</td>
<td>40%</td>
<td>108 $</td>
</tr>
<tr>
<td>CKH</td>
<td>24%</td>
<td>58%</td>
<td>-1248 $</td>
</tr>
<tr>
<td>CMCSA</td>
<td>46%</td>
<td>22%</td>
<td>25599 $</td>
</tr>
<tr>
<td>CNQR</td>
<td>13%</td>
<td>57%</td>
<td>5837 $</td>
</tr>
<tr>
<td>COO</td>
<td>33%</td>
<td>53%</td>
<td>1332 $</td>
</tr>
<tr>
<td>COST</td>
<td>7%</td>
<td>51%</td>
<td>-32206 $</td>
</tr>
<tr>
<td>CPSI</td>
<td>44%</td>
<td>76%</td>
<td>540 $</td>
</tr>
<tr>
<td>CPWR</td>
<td>256%</td>
<td>23%</td>
<td>403 $</td>
</tr>
<tr>
<td>CR</td>
<td>32%</td>
<td>70%</td>
<td>-1619 $</td>
</tr>
<tr>
<td>CRI</td>
<td>101%</td>
<td>76%</td>
<td>-302 $</td>
</tr>
<tr>
<td>CSCO</td>
<td>193%</td>
<td>9%</td>
<td>15086 $</td>
</tr>
<tr>
<td>CSE</td>
<td>294%</td>
<td>27%</td>
<td>220 $</td>
</tr>
<tr>
<td>CSL</td>
<td>128%</td>
<td>60%</td>
<td>-337 $</td>
</tr>
<tr>
<td>CTRN</td>
<td>106%</td>
<td>38%</td>
<td>162 $</td>
</tr>
<tr>
<td>CTSH</td>
<td>18%</td>
<td>51%</td>
<td>16447 $</td>
</tr>
<tr>
<td>DCOM</td>
<td>54%</td>
<td>71%</td>
<td>206 $</td>
</tr>
<tr>
<td>Stock</td>
<td>Change %</td>
<td>% Additional</td>
<td>Profit ($)</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>DELL</td>
<td>133%</td>
<td>7%</td>
<td>-7695 $</td>
</tr>
<tr>
<td>DIS</td>
<td>25%</td>
<td>38%</td>
<td>21363 $</td>
</tr>
<tr>
<td>DK</td>
<td>28%</td>
<td>65%</td>
<td>2580 $</td>
</tr>
<tr>
<td>DOW</td>
<td>394%</td>
<td>35%</td>
<td>921 $</td>
</tr>
<tr>
<td>EBAY</td>
<td>406%</td>
<td>34%</td>
<td>6854 $</td>
</tr>
<tr>
<td>ESRX</td>
<td>10%</td>
<td>47%</td>
<td>34609 $</td>
</tr>
<tr>
<td>EWBC</td>
<td>9%</td>
<td>36%</td>
<td>31326 $</td>
</tr>
<tr>
<td>FCN</td>
<td>122%</td>
<td>65%</td>
<td>101 $</td>
</tr>
<tr>
<td>FFIC</td>
<td>48%</td>
<td>25%</td>
<td>418 $</td>
</tr>
<tr>
<td>FL</td>
<td>11%</td>
<td>50%</td>
<td>-9114 $</td>
</tr>
<tr>
<td>FMER</td>
<td>35%</td>
<td>41%</td>
<td>-4522 $</td>
</tr>
<tr>
<td>FPO</td>
<td>56%</td>
<td>114%</td>
<td>-85 $</td>
</tr>
<tr>
<td>FRED</td>
<td>310%</td>
<td>51%</td>
<td>85 $</td>
</tr>
<tr>
<td>FULT</td>
<td>41%</td>
<td>25%</td>
<td>-4251 $</td>
</tr>
<tr>
<td>GAS</td>
<td>65%</td>
<td>62%</td>
<td>325 $</td>
</tr>
<tr>
<td>GE</td>
<td>382%</td>
<td>11%</td>
<td>-4446 $</td>
</tr>
<tr>
<td>GILD</td>
<td>14%</td>
<td>39%</td>
<td>-57605 $</td>
</tr>
<tr>
<td>GLW</td>
<td>92%</td>
<td>6%</td>
<td>-12881 $</td>
</tr>
<tr>
<td>GOOG</td>
<td>22%</td>
<td>32%</td>
<td>-211247 $</td>
</tr>
<tr>
<td>GPS</td>
<td>61%</td>
<td>48%</td>
<td>3788 $</td>
</tr>
<tr>
<td>HON</td>
<td>10%</td>
<td>70%</td>
<td>13206 $</td>
</tr>
<tr>
<td>HPQ</td>
<td>73%</td>
<td>19%</td>
<td>-18421 $</td>
</tr>
<tr>
<td>IMGN</td>
<td>35%</td>
<td>32%</td>
<td>-5303 $</td>
</tr>
<tr>
<td>INTC</td>
<td>61%</td>
<td>5%</td>
<td>111422 $</td>
</tr>
<tr>
<td>IPAR</td>
<td>57%</td>
<td>22%</td>
<td>-316 $</td>
</tr>
<tr>
<td>ISIL</td>
<td>14680%</td>
<td>28%</td>
<td>-8 $</td>
</tr>
<tr>
<td>ISRG</td>
<td>320%</td>
<td>45%</td>
<td>-6180 $</td>
</tr>
<tr>
<td>Ticker</td>
<td>19%</td>
<td>56%</td>
<td>-2189 $</td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td>-----</td>
<td>----------</td>
</tr>
<tr>
<td>JKHY</td>
<td>0%</td>
<td>74%</td>
<td>26968 $</td>
</tr>
<tr>
<td>KMB</td>
<td>310%</td>
<td>36%</td>
<td>-839 $</td>
</tr>
<tr>
<td>KR</td>
<td>110%</td>
<td>33%</td>
<td>466 $</td>
</tr>
<tr>
<td>LANC</td>
<td>213%</td>
<td>60%</td>
<td>273 $</td>
</tr>
<tr>
<td>LECO</td>
<td>56%</td>
<td>47%</td>
<td>2514 $</td>
</tr>
<tr>
<td>LPNT</td>
<td>357%</td>
<td>57%</td>
<td>272 $</td>
</tr>
<tr>
<td>LSTR</td>
<td>42%</td>
<td>41%</td>
<td>2031 $</td>
</tr>
<tr>
<td>MAKO</td>
<td>53%</td>
<td>60%</td>
<td>507 $</td>
</tr>
<tr>
<td>MANT</td>
<td>8%</td>
<td>56%</td>
<td>8883 $</td>
</tr>
<tr>
<td>MDCO</td>
<td>17%</td>
<td>36%</td>
<td>14901 $</td>
</tr>
<tr>
<td>MELI</td>
<td>62%</td>
<td>26%</td>
<td>354 $</td>
</tr>
<tr>
<td>MIG</td>
<td>66%</td>
<td>47%</td>
<td>-4133 $</td>
</tr>
<tr>
<td>MMM</td>
<td>61%</td>
<td>75%</td>
<td>-41 $</td>
</tr>
<tr>
<td>MOD</td>
<td>8%</td>
<td>60%</td>
<td>-10969 $</td>
</tr>
<tr>
<td>MOS</td>
<td>45%</td>
<td>42%</td>
<td>-845 $</td>
</tr>
<tr>
<td>MRTN</td>
<td>191276%</td>
<td>63%</td>
<td>0 $</td>
</tr>
<tr>
<td>MXWL</td>
<td>33%</td>
<td>67%</td>
<td>975 $</td>
</tr>
<tr>
<td>NUS</td>
<td>10%</td>
<td>42%</td>
<td>22755 $</td>
</tr>
<tr>
<td>NXTM</td>
<td>275%</td>
<td>40%</td>
<td>76 $</td>
</tr>
<tr>
<td>PBH</td>
<td>27%</td>
<td>43%</td>
<td>436 $</td>
</tr>
<tr>
<td>PFE</td>
<td>51%</td>
<td>15%</td>
<td>26718 $</td>
</tr>
<tr>
<td>PG</td>
<td>2356%</td>
<td>34%</td>
<td>337 $</td>
</tr>
<tr>
<td>PNC</td>
<td>25%</td>
<td>65%</td>
<td>-6350 $</td>
</tr>
<tr>
<td>PNY</td>
<td>16%</td>
<td>58%</td>
<td>1485 $</td>
</tr>
<tr>
<td>PTP</td>
<td>47%</td>
<td>69%</td>
<td>803 $</td>
</tr>
<tr>
<td>RIGL</td>
<td>54%</td>
<td>29%</td>
<td>-2456 $</td>
</tr>
<tr>
<td>RO</td>
<td>ROC</td>
<td>27%</td>
<td>58%</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>RO</td>
<td>ROCK</td>
<td>162%</td>
<td>38%</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td>8%</td>
<td>84%</td>
</tr>
<tr>
<td>SF</td>
<td>SFG</td>
<td>28%</td>
<td>57%</td>
</tr>
<tr>
<td>SW</td>
<td>SWN</td>
<td>37%</td>
<td>45%</td>
</tr>
</tbody>
</table>

Relative mispricing of wealth using the frictionless wealth equation, ratio between adverse selection and transaction cost friction terms and end-of-day gains of the representative liquidity provider on 04/18/13.

### 3.3 Price impact and adverse selection

For our empirical analysis of adverse selection, we emphasize the two cases treated: with and without reconstruction of parent trades. The theory remains the same in both cases, but results are stronger when parent orders are reconstructed, as there is rarely price impact in between child orders of the same parent trade.

#### 3.3.1 Adverse selection

In its simplest form, we model adverse selection with the inequality

\[ \Delta_n L \Delta_n p \leq 0 \]  \hspace{1cm} (3.13)

given that the trader we consider only uses limit orders.

**Informed trading**

A straightforward way to quantify this relationship is therefore to compute the proportion of trades such that
1. $\Delta_n L \Delta_n p < 0$;
2. $\Delta_n L \Delta_n p = 0$;
3. $\Delta_n L \Delta_n p > 0$.

Case 1 means that the trade led to price impact and was most likely adversely selected. Case 2 means that the trade left the midprice unchanged and may be considered as 'noise' trading. Case 3 corresponds to reverse price impact, which should happen rarely when adverse selection is present.

Table

For most stocks, the number of trades in case 3 is relatively small and we are confident validating our model. For the stocks where the proportion is relatively high, the number of trades in case 1 is an order of magnitude larger.

The resulting table of proportions of trades in each of the three cases does not amount to a true econometric test. It serves as a guiding statistics to build up intuition and formulate our more sophisticated continuous time model and the corresponding hypothesis to test. The next subsection continues in this line of thought of building up intuition before the actual econometric test is presented and executed.

<table>
<thead>
<tr>
<th>stock</th>
<th>total trades</th>
<th>with impact</th>
<th>without impact</th>
<th>reverse impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>5115</td>
<td>325 (6.35%)</td>
<td>4789</td>
<td>1 (0.02%)</td>
</tr>
<tr>
<td>AAPL</td>
<td>46709</td>
<td>16896 (36.17%)</td>
<td>27466</td>
<td>2347 (5.03%)</td>
</tr>
<tr>
<td>ADBE</td>
<td>9031</td>
<td>2270 (25.13%)</td>
<td>6613</td>
<td>148 (1.64%)</td>
</tr>
<tr>
<td>AGN</td>
<td>1884</td>
<td>1073 (56.95%)</td>
<td>656</td>
<td>155 (8.23%)</td>
</tr>
<tr>
<td>AINV</td>
<td>1882</td>
<td>199 (10.57%)</td>
<td>1678</td>
<td>5 (0.27%)</td>
</tr>
<tr>
<td>AMAT</td>
<td>9383</td>
<td>524 (5.58%)</td>
<td>8854</td>
<td>5 (0.06%)</td>
</tr>
<tr>
<td>AMED</td>
<td>697</td>
<td>195 (27.97%)</td>
<td>478</td>
<td>24 (3.45%)</td>
</tr>
<tr>
<td>AMGN</td>
<td>15200</td>
<td>6494 (42.72%)</td>
<td>8161</td>
<td>545 (3.59%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>AMZN</td>
<td>11966</td>
<td>6098 (50.96%)</td>
<td>5219</td>
<td>649 (5.43%)</td>
</tr>
<tr>
<td>ANGO</td>
<td>482</td>
<td>128 (26.55%)</td>
<td>337</td>
<td>17 (3.53%)</td>
</tr>
<tr>
<td>APOG</td>
<td>968</td>
<td>411 (42.45%)</td>
<td>501</td>
<td>56 (5.79%)</td>
</tr>
<tr>
<td>ARCC</td>
<td>2242</td>
<td>246 (10.97%)</td>
<td>1992</td>
<td>4 (0.18%)</td>
</tr>
<tr>
<td>AXP</td>
<td>14395</td>
<td>4070 (28.27%)</td>
<td>9982</td>
<td>343 (2.39%)</td>
</tr>
<tr>
<td>AYI</td>
<td>732</td>
<td>438 (59.83%)</td>
<td>231</td>
<td>63 (8.61%)</td>
</tr>
<tr>
<td>AZZ</td>
<td>102</td>
<td>63 (61.76%)</td>
<td>21</td>
<td>18 (17.65%)</td>
</tr>
<tr>
<td>BAS</td>
<td>469</td>
<td>268 (57.14%)</td>
<td>163</td>
<td>38 (8.11%)</td>
</tr>
<tr>
<td>BHI</td>
<td>9648</td>
<td>3474 (36%)</td>
<td>5818</td>
<td>356 (3.69%)</td>
</tr>
<tr>
<td>BIIB</td>
<td>4270</td>
<td>2287 (53.55%)</td>
<td>1658</td>
<td>325 (7.62%)</td>
</tr>
<tr>
<td>BRCM</td>
<td>18273</td>
<td>3310 (18.11%)</td>
<td>14839</td>
<td>124 (0.68%)</td>
</tr>
<tr>
<td>BRE</td>
<td>1141</td>
<td>490 (42.94%)</td>
<td>580</td>
<td>71 (6.23%)</td>
</tr>
<tr>
<td>BXS</td>
<td>672</td>
<td>272 (40.47%)</td>
<td>377</td>
<td>23 (3.43%)</td>
</tr>
<tr>
<td>BZ</td>
<td>299</td>
<td>110 (36.78%)</td>
<td>167</td>
<td>22 (7.36%)</td>
</tr>
<tr>
<td>CB</td>
<td>1906</td>
<td>924 (48.47%)</td>
<td>879</td>
<td>103 (5.41%)</td>
</tr>
<tr>
<td>CBEY</td>
<td>519</td>
<td>168 (32.36%)</td>
<td>315</td>
<td>36 (6.94%)</td>
</tr>
<tr>
<td>CBT</td>
<td>645</td>
<td>334 (51.78%)</td>
<td>242</td>
<td>69 (10.7%)</td>
</tr>
<tr>
<td>CBZ</td>
<td>121</td>
<td>47 (38.84%)</td>
<td>66</td>
<td>8 (6.62%)</td>
</tr>
<tr>
<td>CDR</td>
<td>166</td>
<td>58 (34.93%)</td>
<td>104</td>
<td>4 (2.41%)</td>
</tr>
<tr>
<td>CELG</td>
<td>7699</td>
<td>3828 (49.72%)</td>
<td>3339</td>
<td>532 (6.91%)</td>
</tr>
<tr>
<td>CETV</td>
<td>248</td>
<td>72 (29.03%)</td>
<td>166</td>
<td>10 (4.04%)</td>
</tr>
<tr>
<td>CKH</td>
<td>212</td>
<td>123 (58.01%)</td>
<td>63</td>
<td>26 (12.27%)</td>
</tr>
<tr>
<td>CMCSA</td>
<td>17870</td>
<td>2410 (13.48%)</td>
<td>15424</td>
<td>36 (0.21%)</td>
</tr>
<tr>
<td>CNQR</td>
<td>976</td>
<td>467 (47.84%)</td>
<td>443</td>
<td>66 (6.77%)</td>
</tr>
<tr>
<td>COO</td>
<td>221</td>
<td>127 (57.46%)</td>
<td>68</td>
<td>26 (11.77%)</td>
</tr>
<tr>
<td>COST</td>
<td>8229</td>
<td>3578 (43.48%)</td>
<td>4346</td>
<td>305 (3.71%)</td>
</tr>
<tr>
<td>CPSI</td>
<td>334</td>
<td>167 (50%)</td>
<td>136</td>
<td>31 (9.29%)</td>
</tr>
<tr>
<td>Symbol</td>
<td>Shares</td>
<td>Volume</td>
<td>Percentage</td>
<td>Shares</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>------------</td>
<td>--------</td>
</tr>
<tr>
<td>CPWR</td>
<td>1855</td>
<td>273</td>
<td>(14.71%)</td>
<td>1576</td>
</tr>
<tr>
<td>CR</td>
<td>590</td>
<td>302</td>
<td>(51.18%)</td>
<td>238</td>
</tr>
<tr>
<td>CRI</td>
<td>925</td>
<td>487</td>
<td>(52.64%)</td>
<td>369</td>
</tr>
<tr>
<td>CSCO</td>
<td>20155</td>
<td>1315</td>
<td>(6.52%)</td>
<td>18836</td>
</tr>
<tr>
<td>CSE</td>
<td>1213</td>
<td>169</td>
<td>(13.93%)</td>
<td>1041</td>
</tr>
<tr>
<td>CSL</td>
<td>377</td>
<td>189</td>
<td>(50.13%)</td>
<td>145</td>
</tr>
<tr>
<td>CTRN</td>
<td>141</td>
<td>62</td>
<td>(43.97%)</td>
<td>68</td>
</tr>
<tr>
<td>CTSH</td>
<td>7991</td>
<td>2958</td>
<td>(37.01%)</td>
<td>4821</td>
</tr>
<tr>
<td>DCOM</td>
<td>312</td>
<td>132</td>
<td>(42.3%)</td>
<td>148</td>
</tr>
<tr>
<td>DELL</td>
<td>3741</td>
<td>198</td>
<td>(5.29%)</td>
<td>3542</td>
</tr>
<tr>
<td>DIS</td>
<td>12956</td>
<td>2924</td>
<td>(22.56%)</td>
<td>9895</td>
</tr>
<tr>
<td>DK</td>
<td>1045</td>
<td>549</td>
<td>(52.53%)</td>
<td>381</td>
</tr>
<tr>
<td>DOW</td>
<td>8823</td>
<td>1682</td>
<td>(19.06%)</td>
<td>7075</td>
</tr>
<tr>
<td>EBAY</td>
<td>47060</td>
<td>11716</td>
<td>(24.89%)</td>
<td>34571</td>
</tr>
<tr>
<td>ESRX</td>
<td>10745</td>
<td>2985</td>
<td>(27.78%)</td>
<td>7543</td>
</tr>
<tr>
<td>EWBC</td>
<td>6391</td>
<td>1335</td>
<td>(20.88%)</td>
<td>4983</td>
</tr>
<tr>
<td>FCN</td>
<td>248</td>
<td>135</td>
<td>(54.43%)</td>
<td>98</td>
</tr>
<tr>
<td>FFIC</td>
<td>200</td>
<td>79</td>
<td>(39.5%)</td>
<td>104</td>
</tr>
<tr>
<td>FL</td>
<td>3877</td>
<td>1103</td>
<td>(28.44%)</td>
<td>2716</td>
</tr>
<tr>
<td>FMER</td>
<td>3747</td>
<td>751</td>
<td>(20.04%)</td>
<td>2965</td>
</tr>
<tr>
<td>FPO</td>
<td>256</td>
<td>146</td>
<td>(57.03%)</td>
<td>97</td>
</tr>
<tr>
<td>FRED</td>
<td>751</td>
<td>283</td>
<td>(37.68%)</td>
<td>435</td>
</tr>
<tr>
<td>FULT</td>
<td>3202</td>
<td>455</td>
<td>(14.2%)</td>
<td>2736</td>
</tr>
<tr>
<td>GAS</td>
<td>643</td>
<td>306</td>
<td>(47.58%)</td>
<td>277</td>
</tr>
<tr>
<td>GE</td>
<td>12968</td>
<td>916</td>
<td>(7.06%)</td>
<td>12048</td>
</tr>
<tr>
<td>GILD</td>
<td>20404</td>
<td>4745</td>
<td>(23.25%)</td>
<td>15370</td>
</tr>
<tr>
<td>GLW</td>
<td>7346</td>
<td>403</td>
<td>(5.48%)</td>
<td>6942</td>
</tr>
<tr>
<td>Symbol</td>
<td>Shares</td>
<td>Percentage</td>
<td>Market Cap</td>
<td>PE Ratio</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>------------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>GOOG</td>
<td>8594</td>
<td>3919 (45.6%)</td>
<td>4066</td>
<td>609 (7.09%)</td>
</tr>
<tr>
<td>GPS</td>
<td>7377</td>
<td>1815 (24.6%)</td>
<td>5465</td>
<td>97 (1.32%)</td>
</tr>
<tr>
<td>HON</td>
<td>5813</td>
<td>2170 (37.33%)</td>
<td>3460</td>
<td>183 (3.15%)</td>
</tr>
<tr>
<td>HPQ</td>
<td>14918</td>
<td>1823 (12.22%)</td>
<td>13076</td>
<td>19 (0.13%)</td>
</tr>
<tr>
<td>IMGN</td>
<td>2416</td>
<td>708 (29.3%)</td>
<td>1627</td>
<td>81 (3.36%)</td>
</tr>
<tr>
<td>INTC</td>
<td>28221</td>
<td>1192 (4.22%)</td>
<td>27022</td>
<td>7 (0.03%)</td>
</tr>
<tr>
<td>IPAR</td>
<td>172</td>
<td>66 (38.37%)</td>
<td>83</td>
<td>23 (13.38%)</td>
</tr>
<tr>
<td>ISIL</td>
<td>2329</td>
<td>382 (16.4%)</td>
<td>1938</td>
<td>9 (0.39%)</td>
</tr>
<tr>
<td>ISRG</td>
<td>2415</td>
<td>1189 (49.23%)</td>
<td>1026</td>
<td>200 (8.29%)</td>
</tr>
<tr>
<td>JKHY</td>
<td>1002</td>
<td>428 (42.71%)</td>
<td>515</td>
<td>59 (5.89%)</td>
</tr>
<tr>
<td>KMB</td>
<td>2608</td>
<td>1414 (54.21%)</td>
<td>981</td>
<td>213 (8.17%)</td>
</tr>
<tr>
<td>KR</td>
<td>6546</td>
<td>1253 (19.14%)</td>
<td>5263</td>
<td>30 (0.46%)</td>
</tr>
<tr>
<td>LANC</td>
<td>226</td>
<td>100 (44.24%)</td>
<td>105</td>
<td>21 (9.3%)</td>
</tr>
<tr>
<td>LECO</td>
<td>1093</td>
<td>526 (48.12%)</td>
<td>503</td>
<td>64 (5.86%)</td>
</tr>
<tr>
<td>LPNT</td>
<td>1876</td>
<td>832 (44.34%)</td>
<td>931</td>
<td>113 (6.03%)</td>
</tr>
<tr>
<td>LSTR</td>
<td>1156</td>
<td>574 (49.65%)</td>
<td>493</td>
<td>89 (7.7%)</td>
</tr>
<tr>
<td>MAKO</td>
<td>1286</td>
<td>339 (26.36%)</td>
<td>920</td>
<td>27 (2.1%)</td>
</tr>
<tr>
<td>MANT</td>
<td>482</td>
<td>201 (41.7%)</td>
<td>242</td>
<td>39 (8.1%)</td>
</tr>
<tr>
<td>MDCO</td>
<td>1737</td>
<td>753 (43.35%)</td>
<td>900</td>
<td>84 (4.84%)</td>
</tr>
<tr>
<td>MELI</td>
<td>812</td>
<td>361 (44.45%)</td>
<td>380</td>
<td>71 (8.75%)</td>
</tr>
<tr>
<td>MIG</td>
<td>528</td>
<td>71 (13.44%)</td>
<td>451</td>
<td>6 (1.14%)</td>
</tr>
<tr>
<td>MMM</td>
<td>5036</td>
<td>2074 (41.18%)</td>
<td>2682</td>
<td>280 (5.56%)</td>
</tr>
<tr>
<td>MOD</td>
<td>110</td>
<td>47 (42.72%)</td>
<td>56</td>
<td>7 (6.37%)</td>
</tr>
<tr>
<td>MOS</td>
<td>3754</td>
<td>1549 (41.26%)</td>
<td>2016</td>
<td>189 (5.04%)</td>
</tr>
<tr>
<td>MRTN</td>
<td>378</td>
<td>141 (37.3%)</td>
<td>212</td>
<td>25 (6.62%)</td>
</tr>
<tr>
<td>MXWL</td>
<td>496</td>
<td>175 (35.28%)</td>
<td>313</td>
<td>8 (1.62%)</td>
</tr>
<tr>
<td>NSR</td>
<td>556</td>
<td>258 (46.4%)</td>
<td>247</td>
<td>51 (9.18%)</td>
</tr>
<tr>
<td>stock</td>
<td>total trades</td>
<td>with impact</td>
<td>without impact</td>
<td>reverse impact</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>-------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>NUS</td>
<td>1300</td>
<td>629 (48.38%)</td>
<td>561</td>
<td>110 (8.47%)</td>
</tr>
<tr>
<td>NXTM</td>
<td>523</td>
<td>140 (26.76%)</td>
<td>359</td>
<td>24 (4.59%)</td>
</tr>
<tr>
<td>PBH</td>
<td>123</td>
<td>60 (48.78%)</td>
<td>43</td>
<td>20 (16.27%)</td>
</tr>
<tr>
<td>PFE</td>
<td>14626</td>
<td>1314 (8.98%)</td>
<td>13303</td>
<td>9 (0.07%)</td>
</tr>
<tr>
<td>PG</td>
<td>18615</td>
<td>3768 (20.24%)</td>
<td>14658</td>
<td>189 (1.02%)</td>
</tr>
<tr>
<td>PNC</td>
<td>5277</td>
<td>2015 (38.18%)</td>
<td>3075</td>
<td>187 (3.55%)</td>
</tr>
<tr>
<td>PNY</td>
<td>540</td>
<td>248 (45.92%)</td>
<td>231</td>
<td>61 (11.3%)</td>
</tr>
<tr>
<td>PTP</td>
<td>513</td>
<td>294 (57.3%)</td>
<td>172</td>
<td>47 (9.17%)</td>
</tr>
<tr>
<td>RIGL</td>
<td>2079</td>
<td>371 (17.84%)</td>
<td>1693</td>
<td>15 (0.73%)</td>
</tr>
<tr>
<td>ROC</td>
<td>968</td>
<td>523 (54.02%)</td>
<td>336</td>
<td>109 (11.27%)</td>
</tr>
<tr>
<td>ROCK</td>
<td>743</td>
<td>214 (28.8%)</td>
<td>507</td>
<td>22 (2.97%)</td>
</tr>
<tr>
<td>SF</td>
<td>440</td>
<td>242 (55%)</td>
<td>148</td>
<td>50 (11.37%)</td>
</tr>
<tr>
<td>SFG</td>
<td>170</td>
<td>98 (57.64%)</td>
<td>48</td>
<td>24 (14.12%)</td>
</tr>
<tr>
<td>SWN</td>
<td>13815</td>
<td>3635 (26.31%)</td>
<td>9953</td>
<td>227 (1.65%)</td>
</tr>
</tbody>
</table>

Absolute number of trades and percentages for each of the three cases for 102 stocks on 04/18/13. Without reconstruction of parent trades.
<table>
<thead>
<tr>
<th>Ticker</th>
<th>Shares</th>
<th>Float</th>
<th>Avg</th>
<th>Last</th>
<th>Vol</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANGO</td>
<td>341</td>
<td>128 (37.53%)</td>
<td>196</td>
<td>17 (4.99%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOG</td>
<td>675</td>
<td>407 (60.29%)</td>
<td>213</td>
<td>55 (8.15%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARCC</td>
<td>926</td>
<td>246 (26.56%)</td>
<td>676</td>
<td>4 (0.44%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AXP</td>
<td>8147</td>
<td>3985 (48.91%)</td>
<td>3814</td>
<td>348 (4.28%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AYI</td>
<td>663</td>
<td>437 (65.91%)</td>
<td>163</td>
<td>63 (9.51%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AZZ</td>
<td>102</td>
<td>63 (61.76%)</td>
<td>21</td>
<td>18 (17.65%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAS</td>
<td>393</td>
<td>268 (68.19%)</td>
<td>87</td>
<td>38 (9.67%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BHI</td>
<td>5966</td>
<td>3444 (57.72%)</td>
<td>2165</td>
<td>357 (5.99%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIIB</td>
<td>3633</td>
<td>2256 (62.09%)</td>
<td>1055</td>
<td>322 (8.87%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRCM</td>
<td>6849</td>
<td>3269 (47.72%)</td>
<td>3452</td>
<td>128 (1.87%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BRE</td>
<td>896</td>
<td>486 (54.24%)</td>
<td>340</td>
<td>70 (7.82%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BXS</td>
<td>495</td>
<td>272 (54.94%)</td>
<td>200</td>
<td>23 (4.65%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BZ</td>
<td>183</td>
<td>110 (60.1%)</td>
<td>51</td>
<td>22 (12.03%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CB</td>
<td>1393</td>
<td>919 (65.97%)</td>
<td>371</td>
<td>103 (7.4%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBYEY</td>
<td>404</td>
<td>168 (41.58%)</td>
<td>200</td>
<td>36 (8.92%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBT</td>
<td>527</td>
<td>327 (62.04%)</td>
<td>132</td>
<td>68 (12.91%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBZ</td>
<td>100</td>
<td>45 (45%)</td>
<td>47</td>
<td>8 (8%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDR</td>
<td>113</td>
<td>58 (51.32%)</td>
<td>51</td>
<td>4 (3.54%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CELG</td>
<td>5602</td>
<td>3737 (66.7%)</td>
<td>1332</td>
<td>533 (9.52%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CETV</td>
<td>157</td>
<td>72 (45.85%)</td>
<td>75</td>
<td>10 (6.37%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CKH</td>
<td>192</td>
<td>122 (63.54%)</td>
<td>44</td>
<td>26 (13.55%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMCSA</td>
<td>6410</td>
<td>2402 (37.47%)</td>
<td>3969</td>
<td>39 (0.61%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNQR</td>
<td>769</td>
<td>465 (60.46%)</td>
<td>238</td>
<td>66 (8.59%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COO</td>
<td>202</td>
<td>125 (61.88%)</td>
<td>51</td>
<td>26 (12.88%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COST</td>
<td>5473</td>
<td>3483 (63.63%)</td>
<td>1687</td>
<td>303 (5.54%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPSI</td>
<td>297</td>
<td>164 (55.21%)</td>
<td>102</td>
<td>31 (10.44%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPWR</td>
<td>686</td>
<td>273 (39.79%)</td>
<td>407</td>
<td>6 (0.88%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symbol</td>
<td>Volume</td>
<td>Shares</td>
<td>Average</td>
<td>Low</td>
<td>Volume</td>
<td>Shares</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>---------</td>
<td>-----</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>CR</td>
<td>493</td>
<td>298</td>
<td>(60.44%)</td>
<td>146</td>
<td>49</td>
<td>(9.94%)</td>
</tr>
<tr>
<td>CRI</td>
<td>703</td>
<td>483</td>
<td>(68.7%)</td>
<td>151</td>
<td>69</td>
<td>(9.82%)</td>
</tr>
<tr>
<td>CSCO</td>
<td>5025</td>
<td>1315</td>
<td>(26.16%)</td>
<td>3706</td>
<td>4</td>
<td>(0.08%)</td>
</tr>
<tr>
<td>CSE</td>
<td>486</td>
<td>168</td>
<td>(34.56%)</td>
<td>316</td>
<td>2</td>
<td>(0.42%)</td>
</tr>
<tr>
<td>CSL</td>
<td>338</td>
<td>187</td>
<td>(55.32%)</td>
<td>108</td>
<td>43</td>
<td>(12.73%)</td>
</tr>
<tr>
<td>CTRN</td>
<td>112</td>
<td>62</td>
<td>(55.35%)</td>
<td>39</td>
<td>11</td>
<td>(9.83%)</td>
</tr>
<tr>
<td>CTSH</td>
<td>4924</td>
<td>2941</td>
<td>(59.72%)</td>
<td>1770</td>
<td>213</td>
<td>(4.33%)</td>
</tr>
<tr>
<td>DCOM</td>
<td>248</td>
<td>132</td>
<td>(53.22%)</td>
<td>84</td>
<td>32</td>
<td>(12.91%)</td>
</tr>
<tr>
<td>DELL</td>
<td>1162</td>
<td>198</td>
<td>(17.03%)</td>
<td>963</td>
<td>1</td>
<td>(0.09%)</td>
</tr>
<tr>
<td>DIS</td>
<td>5801</td>
<td>2900</td>
<td>(49.99%)</td>
<td>2762</td>
<td>139</td>
<td>(2.4%)</td>
</tr>
<tr>
<td>DK</td>
<td>868</td>
<td>535</td>
<td>(61.63%)</td>
<td>219</td>
<td>114</td>
<td>(13.14%)</td>
</tr>
<tr>
<td>DOW</td>
<td>3552</td>
<td>1679</td>
<td>(47.26%)</td>
<td>1807</td>
<td>66</td>
<td>(1.86%)</td>
</tr>
<tr>
<td>EBAY</td>
<td>21407</td>
<td>11430</td>
<td>(53.39%)</td>
<td>9191</td>
<td>786</td>
<td>(3.68%)</td>
</tr>
<tr>
<td>ESRX</td>
<td>5771</td>
<td>2957</td>
<td>(51.23%)</td>
<td>2595</td>
<td>219</td>
<td>(3.8%)</td>
</tr>
<tr>
<td>EWBC</td>
<td>3000</td>
<td>1328</td>
<td>(44.26%)</td>
<td>1599</td>
<td>73</td>
<td>(2.44%)</td>
</tr>
<tr>
<td>FCN</td>
<td>202</td>
<td>134</td>
<td>(66.33%)</td>
<td>53</td>
<td>15</td>
<td>(7.43%)</td>
</tr>
<tr>
<td>FFIC</td>
<td>150</td>
<td>75</td>
<td>(50%)</td>
<td>58</td>
<td>17</td>
<td>(11.34%)</td>
</tr>
<tr>
<td>FL</td>
<td>2183</td>
<td>1099</td>
<td>(50.34%)</td>
<td>1026</td>
<td>58</td>
<td>(2.66%)</td>
</tr>
<tr>
<td>FMER</td>
<td>2015</td>
<td>751</td>
<td>(37.27%)</td>
<td>1233</td>
<td>31</td>
<td>(1.54%)</td>
</tr>
<tr>
<td>FPO</td>
<td>207</td>
<td>145</td>
<td>(70.04%)</td>
<td>50</td>
<td>12</td>
<td>(5.8%)</td>
</tr>
<tr>
<td>FRED</td>
<td>530</td>
<td>281</td>
<td>(53.01%)</td>
<td>217</td>
<td>32</td>
<td>(6.04%)</td>
</tr>
<tr>
<td>FULT</td>
<td>1626</td>
<td>454</td>
<td>(27.92%)</td>
<td>1161</td>
<td>11</td>
<td>(0.68%)</td>
</tr>
<tr>
<td>GAS</td>
<td>494</td>
<td>302</td>
<td>(61.13%)</td>
<td>132</td>
<td>60</td>
<td>(12.15%)</td>
</tr>
<tr>
<td>GE</td>
<td>3351</td>
<td>913</td>
<td>(27.24%)</td>
<td>2434</td>
<td>4</td>
<td>(0.12%)</td>
</tr>
<tr>
<td>GILD</td>
<td>9700</td>
<td>4707</td>
<td>(48.52%)</td>
<td>4704</td>
<td>289</td>
<td>(2.98%)</td>
</tr>
<tr>
<td>GLW</td>
<td>1703</td>
<td>399</td>
<td>(23.42%)</td>
<td>1303</td>
<td>1</td>
<td>(0.06%)</td>
</tr>
<tr>
<td>GOOG</td>
<td>6745</td>
<td>3766</td>
<td>(55.83%)</td>
<td>2366</td>
<td>613</td>
<td>(9.09%)</td>
</tr>
<tr>
<td>Symbol</td>
<td>Basis</td>
<td>Market Cap (Percentage)</td>
<td>Total Shares</td>
<td>Percentage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------------------------</td>
<td>--------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPS</td>
<td>3446</td>
<td>1809 (52.49%)</td>
<td>1540</td>
<td>97 (2.82%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HON</td>
<td>3622</td>
<td>2151 (59.38%)</td>
<td>1290</td>
<td>181 (5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HPQ</td>
<td>4337</td>
<td>1819 (41.94%)</td>
<td>2498</td>
<td>20 (0.47%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMGN</td>
<td>1461</td>
<td>705 (48.25%)</td>
<td>674</td>
<td>82 (5.62%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTC</td>
<td>6863</td>
<td>1190 (17.33%)</td>
<td>5666</td>
<td>7 (0.11%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPAR</td>
<td>138</td>
<td>66 (47.82%)</td>
<td>49</td>
<td>23 (16.67%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISIL</td>
<td>1002</td>
<td>381 (38.02%)</td>
<td>612</td>
<td>9 (0.9%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISRG</td>
<td>2015</td>
<td>1165 (57.81%)</td>
<td>654</td>
<td>196 (9.73%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JKH3</td>
<td>721</td>
<td>421 (58.39%)</td>
<td>240</td>
<td>60 (8.33%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KMB</td>
<td>2044</td>
<td>1394 (68.19%)</td>
<td>435</td>
<td>215 (10.52%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KR</td>
<td>2848</td>
<td>1251 (43.92%)</td>
<td>1566</td>
<td>31 (1.09%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LANC</td>
<td>168</td>
<td>99 (58.92%)</td>
<td>48</td>
<td>21 (12.5%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LECO</td>
<td>807</td>
<td>524 (64.93%)</td>
<td>219</td>
<td>64 (7.94%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPNT</td>
<td>1401</td>
<td>821 (58.6%)</td>
<td>465</td>
<td>115 (8.21%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSTR</td>
<td>938</td>
<td>571 (60.87%)</td>
<td>278</td>
<td>89 (9.49%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAKO</td>
<td>756</td>
<td>337 (44.57%)</td>
<td>392</td>
<td>27 (3.58%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MANT</td>
<td>364</td>
<td>201 (55.21%)</td>
<td>124</td>
<td>39 (10.72%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDCO</td>
<td>1219</td>
<td>750 (61.52%)</td>
<td>385</td>
<td>84 (6.9%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MELI</td>
<td>665</td>
<td>354 (53.23%)</td>
<td>240</td>
<td>71 (10.68%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIG</td>
<td>302</td>
<td>71 (23.5%)</td>
<td>226</td>
<td>5 (1.66%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMM</td>
<td>3212</td>
<td>2058 (64.07%)</td>
<td>874</td>
<td>280 (8.72%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOS</td>
<td>2411</td>
<td>1541 (63.91%)</td>
<td>681</td>
<td>189 (7.84%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRTN</td>
<td>307</td>
<td>141 (45.92%)</td>
<td>141</td>
<td>25 (8.15%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MXWL</td>
<td>327</td>
<td>175 (53.51%)</td>
<td>144</td>
<td>8 (2.45%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSR</td>
<td>464</td>
<td>256 (55.17%)</td>
<td>157</td>
<td>51 (11%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUS</td>
<td>1030</td>
<td>625 (60.67%)</td>
<td>295</td>
<td>110 (10.68%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NXTM</td>
<td>369</td>
<td>140 (37.94%)</td>
<td>205</td>
<td>24 (6.51%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Absolute number of trades and percentages for each of the three cases for 103 stocks on 04/18/13. With reconstruction of parent trades.

**Adverse selection loss**

A more graphical way to check for adverse selection is to compute the cumulative losses due to adverse selection. This corresponds to the third term of our wealth equation. Mathematically, this quantity is called the empirical \textit{quadratic covariation} between the midprice and the inventory and -as the name suggests- measures whether the two quantities move in a correlated fashion.

\[
\sum_{n=1}^{N} \Delta_n p \Delta_n L
\]  

(3.14)

If the adverse selection inequality holds, this quadratic covariation would be a decreasing function of time. In a frictionless market, it would follow a random walk.
Figure 3.3: Empirical cumulative losses due to adverse selection (rescaled) for selected stocks for which we do not reject our model. Rejection is based on the econometric test below.

While this is not a rigorous econometric test, it is a very appealing way to visualize the actual cumulative losses due to adverse selection.

The graphs of figure 3.3 are consistent with the test results of section 3.3.1. Even though some of the stocks considered may have a proportion of trades of type (3) in the high teens, the corresponding quadratic covariations are still -on the macroscopic
scale-decreasing, suggesting that these trades are completely drowned among the trades of type [1].

**Remark 3.3.1.** These graphs have an important consequence for continuous time models. Under the assumption that the midprice is an Itô process, the fact that the quadratic covariation between the midprice and the inventory is non-negligible, and decreasing, implies that the inventory is an Itô process, with a negative instantaneous correlation coefficient. This in particular rules out models where the price is assumed to have unbounded variations but not the inventory.

We provide a more rigorous econometric model in the following section.

**A diffusion model**

Finally, our econometric test for adverse selection is a diffusion model for both the midprice and the aggregate liquidity provider’s inventory. Assume that

\[
\begin{align*}
dp_t &= \mu_t dt + \sigma_t dW_t \\
DL_t &= b_t dt + l_t dW'_t
\end{align*}
\]

(3.15)

with instantaneous correlation \(\rho_t\) between \(W_t\) and \(W'_t\). Then the adverse selection model implies that

\[
\rho_t \leq 0
\]

(3.16)

that is, the changes of a liquidity provider’s inventory are always negatively correlated to the price returns.

We therefore want to reject the following null hypothesis:

**Assumption 3.3.2.** There exists \(t \in [0, 1]\) such that \(\rho_t > 0\).

If we then denote by \(p^N\) and \(L^N\) the discrete measurements of the diffusion processes \(p\) and \(L\) on the uniform grid \(\{1/N, 2/N, ..., 1\}\) then [1] tells us to consider the
discrete processes:

\[
\begin{aligned}
C_t &= \sum_{n=1}^{\lceil N_t \rceil - 1} \Delta_n p^N \Delta_n L^N \\
V_t^N &= N \sum_{n=1}^{\lceil N_t \rceil - 2} \left( (\Delta_n p^N \Delta_{n+1} L^N)^2 + \Delta_n p^N \Delta_n L^N \Delta_{n+1} p^N \Delta_{n+1} L^N \right)
\end{aligned}
\]  

(3.17)

which leads to the functional central limit theorem

\[
\mathcal{L} \left( \frac{C_t^N - [p, L]_t}{\sqrt{N^{-1} |V_t^N|}} \right) \to N(0, 1)
\]

(3.18)

this can then be used to compute rejection probabilities of the null hypothesis on buckets of \(M\) trades assuming that \(\rho_t, \sigma_t\) and \(l_t\) are constant on these buckets. We then multiply those rejection probabilities to obtain the overall rejection probability for the null hypothesis.

We chose \(M\) for each stock such that the trading day was divided into 8 buckets.

<table>
<thead>
<tr>
<th>stock</th>
<th>prob rej</th>
<th>stock</th>
<th>prob rej</th>
<th>stock</th>
<th>prob rej</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>0.99997</td>
<td>CPSI</td>
<td>0.9908</td>
<td>IPAR</td>
<td>0.86019</td>
</tr>
<tr>
<td>AAPL</td>
<td>0.99999</td>
<td>CPWR</td>
<td>0.99631</td>
<td>ISIL</td>
<td>0.99981</td>
</tr>
<tr>
<td>ADBE</td>
<td>0.98026</td>
<td>CR</td>
<td>0.98686</td>
<td>ISRG</td>
<td>0.99974</td>
</tr>
<tr>
<td>AGN</td>
<td>0.98824</td>
<td>CRI</td>
<td>0.99576</td>
<td>JKHY</td>
<td>0.99807</td>
</tr>
<tr>
<td>AINV</td>
<td>0.99136</td>
<td>CSCO</td>
<td>0.71215</td>
<td>KMB</td>
<td>0.67986</td>
</tr>
<tr>
<td>AMAT</td>
<td>0.99982</td>
<td>CSE</td>
<td>0.99932</td>
<td>KR</td>
<td>0.99832</td>
</tr>
<tr>
<td>AMED</td>
<td>0.99917</td>
<td>CSL</td>
<td>0.99368</td>
<td>LANC</td>
<td>0.99891</td>
</tr>
<tr>
<td>AMGN</td>
<td>0.99604</td>
<td>CTRN</td>
<td>0.99998</td>
<td>LECO</td>
<td>0.99872</td>
</tr>
<tr>
<td>AMZN</td>
<td>0.99963</td>
<td>CTSN</td>
<td>0.99466</td>
<td>LPNT</td>
<td>0.99139</td>
</tr>
<tr>
<td>ANGO</td>
<td>0.99967</td>
<td>DCOM</td>
<td>0.99915</td>
<td>LSTR</td>
<td>0.99949</td>
</tr>
<tr>
<td>APOG</td>
<td>0.99555</td>
<td>DELL</td>
<td>0.9893</td>
<td>MAKO</td>
<td>0.99517</td>
</tr>
<tr>
<td>ARCC</td>
<td>0.99973</td>
<td>DIS</td>
<td>0.37821</td>
<td>MANT</td>
<td>0.99997</td>
</tr>
<tr>
<td>AXP</td>
<td>0.98384</td>
<td>DK</td>
<td>0.99584</td>
<td>MDCO</td>
<td>0.99998</td>
</tr>
<tr>
<td>stock</td>
<td>prob rej</td>
<td>stock</td>
<td>prob rej</td>
<td>stock</td>
<td>prob rej</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>-------</td>
<td>----------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>AYI</td>
<td>0.97338</td>
<td>DOW</td>
<td>0.99915</td>
<td>MELI</td>
<td>0.99599</td>
</tr>
<tr>
<td>AZZ</td>
<td>0.99754</td>
<td>EBAY</td>
<td>0.99955</td>
<td>MIG</td>
<td>0.99227</td>
</tr>
<tr>
<td>BAS</td>
<td>0.97226</td>
<td>ESRX</td>
<td>0.90279</td>
<td>MMM</td>
<td>0.37671</td>
</tr>
<tr>
<td>BHI</td>
<td>0.53919</td>
<td>EWBC</td>
<td>0.99677</td>
<td>MOD</td>
<td>0.99999</td>
</tr>
<tr>
<td>BIIB</td>
<td>0.96129</td>
<td>FCN</td>
<td>0.99749</td>
<td>MOS</td>
<td>0.81274</td>
</tr>
<tr>
<td>BRCM</td>
<td>0.99937</td>
<td>FFIC</td>
<td>0.99918</td>
<td>MRTN</td>
<td>0.99972</td>
</tr>
<tr>
<td>BRE</td>
<td>0.98179</td>
<td>FL</td>
<td>0.99922</td>
<td>MXWL</td>
<td>0.99689</td>
</tr>
<tr>
<td>BXS</td>
<td>0.99955</td>
<td>FMER</td>
<td>0.99724</td>
<td>NSR</td>
<td>0.91632</td>
</tr>
<tr>
<td>BZ</td>
<td>0.98989</td>
<td>FPO</td>
<td>0.99995</td>
<td>NUS</td>
<td>0.99813</td>
</tr>
<tr>
<td>CB</td>
<td>0.99649</td>
<td>FRED</td>
<td>0.99999</td>
<td>NXTM</td>
<td>0.99999</td>
</tr>
<tr>
<td>CBEY</td>
<td>0.18014</td>
<td>FULT</td>
<td>0.9754</td>
<td>PBH</td>
<td>0.98192</td>
</tr>
<tr>
<td>CBT</td>
<td>0.98376</td>
<td>GAS</td>
<td>0.97304</td>
<td>PFE</td>
<td>0.99991</td>
</tr>
<tr>
<td>CBZ</td>
<td>0.99992</td>
<td>GE</td>
<td>0.99914</td>
<td>PG</td>
<td>0.99985</td>
</tr>
<tr>
<td>CDR</td>
<td>0.99999</td>
<td>GILD</td>
<td>0.99982</td>
<td>PNC</td>
<td>0.99164</td>
</tr>
<tr>
<td>CELG</td>
<td>0.99991</td>
<td>GLW</td>
<td>0.99963</td>
<td>PNY</td>
<td>0.96944</td>
</tr>
<tr>
<td>CETV</td>
<td>0.99364</td>
<td>GOOG</td>
<td>0.99292</td>
<td>PTP</td>
<td>0.96567</td>
</tr>
<tr>
<td>CKH</td>
<td>0.9937</td>
<td>GPS</td>
<td>0.9999</td>
<td>RIGL</td>
<td>0.71092</td>
</tr>
<tr>
<td>CMCSA</td>
<td>0.99573</td>
<td>HON</td>
<td>0.1314</td>
<td>ROC</td>
<td>0.81843</td>
</tr>
<tr>
<td>CNQR</td>
<td>0.96989</td>
<td>HPQ</td>
<td>0.99944</td>
<td>ROCK</td>
<td>0.98525</td>
</tr>
<tr>
<td>COO</td>
<td>0.84609</td>
<td>IMGN</td>
<td>0.98812</td>
<td>SF</td>
<td>0.93587</td>
</tr>
<tr>
<td>COST</td>
<td>0.99758</td>
<td>INTC</td>
<td>0.98443</td>
<td>SFG</td>
<td>0.99929</td>
</tr>
</tbody>
</table>

Probability of rejection for the null hypothesis on 04/18/13 for stocks with more than 100 trades. Without reconstruction of parent orders.

```
<table>
<thead>
<tr>
<th>stock</th>
<th>prob rej</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>0.96728</td>
</tr>
<tr>
<td>CPSI</td>
<td>0.99755</td>
</tr>
<tr>
<td>IPAR</td>
<td>0.91515</td>
</tr>
</tbody>
</table>
```

63
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
<th>Symbol</th>
<th>Value</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAPL</td>
<td>1</td>
<td>CPWR</td>
<td>0.99889</td>
<td>ISIL</td>
<td>0.82535</td>
</tr>
<tr>
<td>ADBE</td>
<td>0.99999</td>
<td>CR</td>
<td>0.99897</td>
<td>ISRG</td>
<td>0.99996</td>
</tr>
<tr>
<td>AGN</td>
<td>0.99931</td>
<td>CRI</td>
<td>0.99625</td>
<td>JKHY</td>
<td>0.99046</td>
</tr>
<tr>
<td>AINV</td>
<td>0.91669</td>
<td>CSCO</td>
<td>0.99999</td>
<td>KMB</td>
<td>0.99999</td>
</tr>
<tr>
<td>AMAT</td>
<td>0.99682</td>
<td>CSE</td>
<td>0.99838</td>
<td>KR</td>
<td>0.99999</td>
</tr>
<tr>
<td>AMED</td>
<td>0.95456</td>
<td>CSL</td>
<td>0.98687</td>
<td>LANC</td>
<td>0.99987</td>
</tr>
<tr>
<td>AMGN</td>
<td>0.99999</td>
<td>CTRN</td>
<td>0.99999</td>
<td>LECO</td>
<td>0.99981</td>
</tr>
<tr>
<td>AMZN</td>
<td>0.99999</td>
<td>CTSH</td>
<td>0.99999</td>
<td>LPNT</td>
<td>0.99999</td>
</tr>
<tr>
<td>ANGO</td>
<td>0.99997</td>
<td>DCOM</td>
<td>0.99999</td>
<td>LSTR</td>
<td>0.99991</td>
</tr>
<tr>
<td>APOG</td>
<td>0.98397</td>
<td>DELL</td>
<td>0.78661</td>
<td>MAKO</td>
<td>0.91648</td>
</tr>
<tr>
<td>ARCC</td>
<td>0.99519</td>
<td>DIS</td>
<td>0.99999</td>
<td>MANT</td>
<td>0.99888</td>
</tr>
<tr>
<td>AXP</td>
<td>0.99999</td>
<td>DK</td>
<td>0.99812</td>
<td>MDCO</td>
<td>0.99998</td>
</tr>
<tr>
<td>AYI</td>
<td>0.99354</td>
<td>DOW</td>
<td>0.99999</td>
<td>MELI</td>
<td>0.90214</td>
</tr>
<tr>
<td>AZZ</td>
<td>0.99754</td>
<td>EBAY</td>
<td>1</td>
<td>MIG</td>
<td>0.98378</td>
</tr>
<tr>
<td>BAS</td>
<td>0.99999</td>
<td>ESRX</td>
<td>1</td>
<td>MMM</td>
<td>0.90883</td>
</tr>
<tr>
<td>BHI</td>
<td>1</td>
<td>EWBC</td>
<td>0.99999</td>
<td>MOS</td>
<td>0.99999</td>
</tr>
<tr>
<td>BIIB</td>
<td>0.99999</td>
<td>FCN</td>
<td>0.99989</td>
<td>MRTN</td>
<td>0.99528</td>
</tr>
<tr>
<td>BRCM</td>
<td>0.99999</td>
<td>FFIC</td>
<td>0.99999</td>
<td>MXWL</td>
<td>0.99998</td>
</tr>
<tr>
<td>BRE</td>
<td>0.99718</td>
<td>FL</td>
<td>0.99999</td>
<td>NSR</td>
<td>0.96399</td>
</tr>
<tr>
<td>BXS</td>
<td>0.9992</td>
<td>FMER</td>
<td>0.90467</td>
<td>NUS</td>
<td>0.9958</td>
</tr>
<tr>
<td>BZ</td>
<td>0.99976</td>
<td>FPO</td>
<td>0.99999</td>
<td>NXTM</td>
<td>0.9997</td>
</tr>
<tr>
<td>CB</td>
<td>0.99999</td>
<td>FRED</td>
<td>0.99988</td>
<td>PBH</td>
<td>0.99717</td>
</tr>
<tr>
<td>CBEY</td>
<td>0.88292</td>
<td>FULT</td>
<td>0.99946</td>
<td>PFE</td>
<td>0.99999</td>
</tr>
<tr>
<td>CBT</td>
<td>0.98185</td>
<td>GAS</td>
<td>0.99863</td>
<td>PG</td>
<td>1</td>
</tr>
<tr>
<td>CBZ</td>
<td>0.99998</td>
<td>GE</td>
<td>0.99999</td>
<td>PNC</td>
<td>0.99999</td>
</tr>
<tr>
<td>CDR</td>
<td>0.99999</td>
<td>GILD</td>
<td>1</td>
<td>PNY</td>
<td>0.99559</td>
</tr>
<tr>
<td>CELG</td>
<td>0.99999</td>
<td>GLW</td>
<td>0.89184</td>
<td>PTP</td>
<td>0.99237</td>
</tr>
<tr>
<td>CETV</td>
<td>0.9999</td>
<td>GOOG</td>
<td>1</td>
<td>RIGL</td>
<td>0.99953</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>CKH</td>
<td>0.99927</td>
<td>GPS</td>
<td>0.99999</td>
<td>ROC</td>
<td>0.99412</td>
</tr>
<tr>
<td>CMCSA</td>
<td>0.99999</td>
<td>HON</td>
<td>0.99999</td>
<td>ROCK</td>
<td>0.99036</td>
</tr>
<tr>
<td>CNQR</td>
<td>0.98594</td>
<td>HPQ</td>
<td>0.99999</td>
<td>SF</td>
<td>0.99888</td>
</tr>
<tr>
<td>COO</td>
<td>0.92067</td>
<td>IMGN</td>
<td>0.98296</td>
<td>SFG</td>
<td>0.99996</td>
</tr>
<tr>
<td>COST</td>
<td>0.99992</td>
<td>INTC</td>
<td>0.99999</td>
<td>SWN</td>
<td>1</td>
</tr>
</tbody>
</table>

Probability of rejection for the null hypothesis on 04/18/13 for stocks with more than 100 trades. With reconstruction of parent orders. One stock dropped from the list because, after the reconstruction of parent orders, the number of trades was insufficient to compute the statistic.

The table shows that our more rudimentary model based on the proportion of trades falling into the three cases is a reasonable guide for most stocks, but may disagree with the more macroscopic picture on particular choices of stocks. Overall, we believe our model -in particular the continuous time version- to hold for a sufficiently high number of stocks to be of interest to practitioners and academics alike.
Chapter 4

Macroscopic limit

In this chapter we implement the main argument of the thesis: deriving macroscopic, continuous-time equations from the microscopic, discrete equations established in chapter 2.

We first present the main technical tools as well as the overall proof strategy in section 4.1. This involves functional law of large numbers and central limit theorems, as our equations can be expressed as non-linear, path-dependent operators on the price and inventory processes. In section 4.2 we apply this procedure for the self-financing portfolio equation as summarized in section 2.1.4. Finally, once the self-financing portfolio equation is derived, section 4.3 derives the continuous-time equivalents of the two models for price impact and adverse selection proposed in section 2.3.

4.1 Tools and strategy

We recall as an example the discrete equation 2.17 derived in chapter 2 for the wealth of a liquidity provider:

\[ \Delta X = L \Delta p + c(-\Delta L) + \Delta p \Delta L \] (4.1)
Given a path of $p$, $L$ and the transaction cost function $c$, it is therefore possible to track the path of the trader’s wealth. The first and last term of the formula are straightforward to ‘translate’ into continuous time. The first one corresponds to a stochastic integral while the last term is a quadratic covariation term. Only the middle term, which is non-linear, is novel. Care must be taken when deriving the continuous limit of this term, as it may blow up if certain scaling properties are not verified.

We first introduce the law of large numbers and central limit theorem needed to deal with such terms, before proposing our set of scaling hypotheses and overall derivation strategy.

4.1.1 Technical tools

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space supporting an 1-dimensional Wiener process $W$ and $Y$ be a 1-dimensional Itô process of the form

$$Y_t = Y_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s$$

(4.2)

where we consider $t \in [0, 1]$.

The first tool we use in this section is the functional law of large number for discretized process by Jacod and Protter [30]. Let $\phi_{\sigma^2}$ denote the density function of the Gaussian distribution with mean 0 and variance $\sigma^2$.

**Theorem 4.1.1** ((7.2.2) from [30]). Let $(t, y) \rightarrow F_t(y)$ be an $\mathcal{F}_t$-adapted random function that is a.s. continuous in $(t, y)$ and verifies the growth condition $F_t(y) \leq Cy^2$ for some constant $C$. Then we have the following convergence u.c.p. as $N \rightarrow \infty$:

$$\frac{1}{N} \sum_{n=1}^{[Nt]} F_{n/N} \left( \sqrt{N} (Y_{(n+1)/N} - Y_{n/N}) \right) \rightarrow \int_0^t \int_0^t F_s(y) \phi_{\sigma^2}(y) dy \, ds$$

67
where \( \sigma_t^2 = \frac{d[Y,Y]_t}{dt} \).

For convenience, we introduce the notation

\[
\Phi_\sigma(F) = \int F(y) \phi_{\sigma^2}(y) dy
\]

for any real \( \sigma \) and function \( F \).

This result will cover the self-financing equation and the reduced form model of adverse selection. For the structural model of price impact however, a functional central limit theorem from the same reference \[30\] must be used. For this central limit theorem, we first summarize the more elaborate hypothesis before giving the result.

**Assumption 4.1.2.** \((H)+(K)\) from \[30\]

Assume that \( b_t \) and \( \sigma_t \) are progressively measurable, \( b_t \) is locally bounded and \( \sigma_t \) is càdlàg.

Let now \( F : \Omega \times [0,1] \times \mathbb{R} \to \mathbb{R} \) be a random, \( \mathcal{F}_t \)-adapted function that is \( C^1 \) in \( y \) and \( C^0 \) in \( (t,y) \). We will shorten the notation to \( y \mapsto F_t(y) \). Define the following assumption.

**Assumption 4.1.3.** \((7.2.1), (10.3.2), (10.3.3), (10.3.4) and (10.3.7)\) from \[30\]

Assume that a.s. for all \( t \), \( F_t \) is an odd function in \( y \).

Furthermore, assume there exists a function \( g : \mathbb{R} \to \mathbb{R} \) with polynomial growth and a real \( \beta > 1/2 \) such that, for all \( \omega \in \Omega \), \( (t,s) \in [0,1]^2 \) and \( y \in \mathbb{R} \):

\[
|F_t(y)| \leq g(y) \\
|F_t'(y)| \leq g(y) \\
|F_t(y) - F_s(y)| \leq g(y)|t-s|^\beta
\]

Let us now state the last result from \[30\] we will use.
Theorem 4.1.4. (10.3.2) from [30] Assume 4.1.2 and 4.1.3. Then there exists a very good filtered extension of the original space such that we have the following stable convergence in law as $N \to \infty$:

$$\frac{1}{\sqrt{N}} \sum_{n=1}^{[Nt]} F_{n/N} \left( \sqrt{N} (X_{(n+1)/N} - X_{n/N}) \right) \to U_t$$

where

$$U_t = \int_0^t b_s \Phi_{\sigma_s} (F'_s) \, ds + \int_0^t \sqrt{\Phi_{\sigma_s} ((F_s)^2)} dW'_s \quad (4.4)$$

with $W'_t$ a $d$-dimensional Wiener process such that

$$[W', W]_t = \int_0^t \frac{\Phi_{\sigma_s} (id F^k_s)}{\sigma_s \sqrt{\Phi_{\sigma_s} (F^k_s)^2}} \, ds$$

where $id$ is the identity function.

4.1.2 Strategy for the limiting procedure

We proceed as follows:

1. We begin with the continuous processes for the inventory $L$, price $p$ and shape function $\gamma$ as our data.

2. By discretizing them, we obtain the data to plug into the discrete relationship of chapter 2, yielding our discrete time self-financing equation.

3. Finally, we take the limit again to obtain the diffusion limits of our discrete output to obtain our continuous-time relationship.

In discrete time, prices are a pure-jump process, and therefore have finite variations. It is common on larger time scales to consider the price as ‘zoomed out’ enough to be approximated by a diffusion process. Mathematically, this corresponds to a van-
ishing tick size. Recall that tick size is typically of the order of magnitude of the cent\(^1\) that is \(10^{-4}\) relative to the typical stock price. Given the relative roughness of the path of inventories when compared to prices, see for example Figure 3.1, it seems reasonable to also expect high-frequency inventories to be modeled by processes with infinite variation. This leaves us with a choice of how to rescale the shape function \(\gamma\). The key is to note that we already assume the price grid tick size to vanish and to note that the bid-ask spread being of the same order of magnitude as the price jumps, also vanishes with the tick size.

Remark 4.1.5. The bid-ask spread is of the same order of magnitude as the price jumps: the tick size. This implies in particular that the bid-ask spread vanishes in absolute terms in the diffusion limit and should therefore be measured in tick-size. The mathematical consequence of this simple comment is that transaction costs do not diverge as the tick size goes to zero, allowing inventories that have infinite variations in the continuous limit.

Mathematical setup

Let \((\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})\) be a filtered probability space supporting two Wiener processes \(W\) and \(W'\) with unspecified correlation structure. We consider a fixed time interval \([0, 1]\) and give ourselves the following \(\mathcal{F}\)-adapted processes for the price and inventory of a provider:

\[
\begin{cases}
    p_t = p_0 + \int_0^t \mu_u du + \int_0^t \sigma_u dW_u \\
    L_t = L_0 + \int_0^t b_u du + \int_0^t l_u dW'_u
\end{cases}
\]

(4.5)

where \(p_0\) and \(L_0\) are \(\mathcal{F}_0\)-measurable elements of \(L^2\) and \(\mu, \sigma, b\) and \(l\) are \(\mathcal{F}\)-adapted and càdlàg processes. Finally, let \(c : \Omega \times [0, 1] \times \mathbb{R} \to \mathbb{R}^d\) be a random, \(\mathcal{F}_t\)-adapted function that is \(C^0\) in \((t, l)\). Assume \(c\) to be a.s. convex for all \(t\), with a minimum at \(c_t(0) = 0\) and such that \(c_t(l) < Cl^2\) for some constant \(C\). We denote by \(\gamma_t\) its

\(^1\)Decibasis point for some exchanges in the foreign exchange market.
Figure 4.1: Renormalization of the model for the diffusion limit. Time is scaled by $1/N$, prices by $1/\sqrt{N}$ and volume by 1 (unchanged). For example, the $y$-axis of $\gamma$ represents cost, that is $[\text{volume}] \cdot [\text{price}]^2$ which scales in $1/N$. The $x$-axis is expressed in prices and is scaled in $1/\sqrt{N}$, leading to the formula $\gamma^N(\cdot) = \gamma(\sqrt{N} \cdot \cdot)/N$.

Legendre transform, which will represent the shape function of the order book as measured in tick size.

Let $1/\sqrt{N}$ be a vanishing tick size. Define the discretized price process as $p^n = p_{n/N}$ and likewise $L^N$.

We propose the following choice of renormalization for the order book.

$$
\gamma_n^N(x) = \frac{1}{N} \gamma_{n/N} \left( \sqrt{N} x \right) 
$$

(4.6)
This in particular implies

$$c_n^N(l) = \frac{1}{N} c_{n/N} \left( \sqrt{Nl} \right)$$  \quad (4.7)

This follows from the fact that $\gamma$ is defined in tick size and needs to be renormalized appropriately in the discrete approximation, where we want $\gamma^N$ to be expressed in absolute terms.

### 4.2 The self-financing portfolio equation

We recall once more the discrete equation 2.17 derived in chapter 2 for the wealth of a liquidity provider, applied on the discretized version of our data to define our discrete wealth process $X_n^N$

$$\Delta_n X^N = L_n^N \Delta_n p^N + c_n^N (-\Delta_n L^N) + \Delta_n p^N \Delta_n L^N$$  \quad (4.8)

The final step is to prove the convergence of $X_{\lfloor Nt \rfloor}^N$ to a continuous Itô process and determine its dynamics.

#### 4.2.1 General case

**Theorem 4.2.1.** The continuous time relationship between provider wealth $X$, inventory $L$, price $p$ and transaction costs $c$ is:

$$dX_t = L_t dp_t + \Phi_{\xi_t}(c_t) dt + d[L,p]_t$$  \quad (4.9)

where $X_t = \lim_{N \to \infty} X_{\lfloor Nt \rfloor}^N$ u.c.p.
Proof. The result to prove is the u.c.p. convergence of

\[
\frac{1}{N} \sum_{n=1}^{\lfloor tN \rfloor} c_{n/N} \left( -\sqrt{N} \Delta_n L^N \right)
\]

(4.10)

to the integral

\[
\int_0^t \Phi_{t_u}(c_u) du
\]

(4.11)

This is a direct application of theorem [4.1.1].

By symmetry, we therefore have the following self-financing equation for liquidity takers:

Corollary 4.2.2. The continuous time relationship between liquidity taker wealth \( X \), inventory \( L \), price \( p \) and transaction costs \( c \) is:

\[
dX_t = L_t dp_t - \Phi_{t_u}(c_t) dt + d[L,p]_t
\]

(4.12)

Bid-ask spread case

In the case where all trades happen at the best bid or ask price, one has \( c_t(l) = \frac{s_t}{2} |l| \)
where \( s_t \) is the bid-ask spread process. This leads to

Corollary 4.2.3. Let \( p \) be the midprice, \( s \) the bid-ask spread and \( L \) be a trader’s inventory process on an order book where all trades happen at the best bid or ask price. Then the trader’s wealth satisfies

\[
dX_t = L_t dp_t \pm \frac{s_t l_t}{\sqrt{2\pi}} dt + d[L,p]_t
\]

(4.13)

where \( \pm \) is + when trading with limit orders and – otherwise.

Note that if we push the sign of \( \rho_t \) onto \( l_t \), making it signed, we can replace \( \pm \) by +, as \( l_t \) will be of the correct sign.
**Time change**

Note that equation (4.9) was proved in a *trade clock*, which means that all the time-related quantities, such as volatility, must be measured per trade time. While this is a positive feature for high frequency models under this clock (e.g. [8, 11]), it is less advantageous for financial problems working under a different clock. For example, pricing an option with a fixed maturity in the calendar clock may be difficult to do directly from equation (1.3). We therefore discuss how our proposed formula behaves under time-changes, with the canonical time-change being the switch to a calendar clock. Another possible time-change is the switch from a trade clock to a volume clock.

**Definition 4.2.4.** We define a good time change to be an $F_t$-adapted stochastic process $\tau_t$ such that $\tau_0 = 0$ and

$$d\tau_t = n_t^2 dt$$  \hspace{1cm} (4.14)

with $n_t$ uniformly bounded away from zero.

We start from:

$$
\begin{align*}
dp_t &= \mu_t dt + \sigma_t dW_t \\
DL_t &= b_t dt + l_t dW'_t \\
DX_t &= L_t dp_t + \Phi_t(c_t) dt + d[L,p]_t
\end{align*}
$$

with $d[L,p]_t \leq 0$, and we study the processes $\tilde{p}_t = p_{\tau_t}$, $\tilde{L}_t = L_{\tau_t}$ and $\tilde{X}_t = X_{\tau_t}$. Note that all the time-changed processes are now adapted with respect to the time-changed filtration $\tilde{F}_t = F_{\tau_t}$. Note also that the processes $\tilde{W}_t = \int_{\tau_t}^{\tau} 1/n_{\tau_u} dW_u$ and $\tilde{W}'_t = \int_{\tau_t}^{\tau} 1/n_{\tau_u} dW'_u$ are $\tilde{F}_t$ Wiener processes.
A simple chain-rule leads to the time-changed dynamics:

\[
\begin{align*}
    d\tilde{p}_t &= \tilde{\mu}_t dt + \tilde{\sigma}_t d\tilde{W}_t \\
    d\tilde{L}_t &= \tilde{b}_t dt + \tilde{l}_t d\tilde{W}_t' \\
    d\tilde{X}_t &= \tilde{\bar{L}}_t d\tilde{p}_t + \Phi_{\tilde{L}_t}(\tilde{c}_t) dt + d[\tilde{L}, \tilde{p}]_t
\end{align*}
\]  

(4.16)

where

\[
\begin{align*}
    \tilde{\mu}_t &= n_t^2 \mu_t; \quad \tilde{b}_t = n_t^2 b_t \\
    \tilde{\sigma}_t &= n_t \sigma_t; \quad \tilde{l}_t = n_t l_t
\end{align*}
\]

which are standard, as well as the more surprising:

\[
\tilde{c}_t(\cdot) = n_t^2 c_t(\cdot/n_t)
\]  

(4.17)

**Remark 4.2.5.** Part of this result is expected: under the modified time clock, drifts and volatility must be measured by the new unit of time instead of by unit of trade, which corresponds to the factors \(n_t^2\) and \(n_t\). However, the unfortunate result is that the transaction cost function must be updated as well. This update follows the same rational as the rescaling argument used earlier as the price grid gets scaled.

**Remark 4.2.6.** Certain models are independent of the time-change. This is for instance the case when all the trades happen at the best bid or ask, and the bid-ask spread \(s_t\) is of the form \(s_t = \lambda \sigma_t\). Such an assumption would follow the conclusion of the empirical paper [49] which suggests a linear relationship between daily average bid-ask spread and daily average volatility per trade. From a theoretical perspective, this model is stable under time change, in the sense that \(\bar{s}_t = \lambda \bar{\sigma}_t\), a desirable property.
Remark 4.2.7. One could have also from the beginning worked under the changed clock and used the law of large numbers with irregular discretization schemes found in [30] to recover the same result.

4.3 Price impact models

In continuous time, price impact must be understood as an *admissibility constraint* between a trader’s inventory and the price quoted by the market. In a frictionless world, any inventory can be obtained regardless of the price dynamics. However, in presence of microstructure, this is not the case anymore. For market orders, trades lead to price changes as the bid or the ask get pushed and market makers incorporate the information contained into the trade. For limit orders, trades often are caused by or predict a change in price. Most of the microstructure model can be thought of as models of this admissibility constraint between an inventory of a trader and the price path given the choice to trade with market or limit orders. These models can be structural or reduced form as well more or less rigid and explicit.

As previously, we will derive these constraints from microscopic models proposed in chapter 2 through our limiting procedure. Once derived, these constraints serve as a definition of admissibility for a given portfolio to be attainable via a certain trading style.

4.3.1 Reduced form model

In the reduced form model, we wish to append the microscopic inequality $\Delta_n L \Delta_n p \leq 0$ when trading with limit orders to our proposed equations. The setup is therefore exactly the same as in section 4.1.2 and the self-financing portfolio equation derived in section 4.2 still holds.

The result is straightforward.
Proposition 4.3.1. Let $L$ be a liquidity provider’s inventory and $p$ the midprice of the asset she is trading. Assume that for each $N$, the discretized processes $L^N$ and $p^N$ verify the inequality

$$\Delta_n L^N \Delta_n p^N \leq 0$$

(4.18)

for all $n$. Then we have that

$$d[L, p]_t \leq 0$$

(4.19)

that is, the quadratic covariation of $p$ and $L$ is non-increasing.

Proof. This follows from the definition of the quadratic covariation. \qed

Remark 4.3.2. Note that the microscopic assumption

$$\Delta_n L^N \Delta_n p^N \leq 0$$

(4.20)

is much stronger than what is needed to obtain the result. A weaker microscopic assumption could be

$$\mathbb{E} \left[ \Delta_n L^N \Delta_n p^N \mid \mathcal{F}^N_n \right] \leq 0$$

(4.21)

This is why our main statistical test in chapter 3 tests directly for decreasing quadratic covariation, as we feel this is the more robust model.

This leads to the following definition of an admissible portfolio in our reduced form framework:

Definition 4.3.3. Admissible limit order trading strategy

Let $p$ be a price process, $L$ an inventory process obtained solely via limit orders and $c$ the corresponding transaction cost process. We define this triplet to be an admissible limit order trading strategy if $[p, L]_t$ is a decreasing process.
This in particular implies that any inventory that has an increasing quadratic covariation with the price must be obtained via market orders. As we will see later in chapter 5, this simple comment has an impact on the delta-hedging of options.

4.3.2 Structural model

In the case of the structural model presented in section 2.3.3 we start with slightly different data, as the aim is to derive $L$ from $p$ or vice-versa. The same general methodology holds however and the self-financing portfolio equation derived in section 4.2 still holds.

Continuous time setup

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space supporting a Wiener process $W_t$. We will fix either an Itô process

$$p_t = p_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$$  \hspace{1cm} (4.22)

for the price or

$$L_t = L_0 + \int_0^t b_s ds + \int_0^t l_s dW_s$$  \hspace{1cm} (4.23)

for the inventory.

In addition to one of these processes, we also fix an order book shape process $\gamma_t$ and denote by $c_t$ the associated transaction cost process.

Assume $L$ (respectively $p$) verifies Assumption 4.1.2 and $c$ (respectively $\gamma$) satisfies Assumption 4.1.3.

Just as previously, we define the discretized processes $L^N_{n} = L_{n/N}$ (respectively $p^N_{n} = p_{n/N}$) and $c^N_{n}(\cdot) = \frac{1}{N}c_{n/N}(\sqrt{N} \cdot)$ (respectively $\gamma^N_{n}(\cdot) = \frac{1}{N}\gamma_{n/N}(\sqrt{N} \cdot)$).
Main result

The main result is a straightforward application of Theorem 4.1.4. If we are given the inventory $L$ and transaction costs $c$ then we have:

**Theorem 4.3.4.** There exists a very good filtered extension of the original space such that we have the stable convergence in law $p_{\lfloor Nt \rfloor}^N \rightarrow p_t$ with

$$dp_t = -\lambda b_t \Phi_{l_t} (c'_t) \, dt + \lambda \sqrt{\Phi_{l_t} ((c'_t)^2)} dW'_t$$

(4.24)

where

$$[W', W]_t = -\int_0^t \frac{\Phi_{l_s} (idc'_s)}{\sqrt{\Phi_{l_s} ((c'_s)^2)}} \, ds.$$  

(4.25)

In particular,

$$d[p, L]_t = -\Phi_{l_t} (idc'_t) \, dt$$

(4.26)

A completely equivalent result is obtained if the price $p$ and order book shape function $\gamma$ are given:

**Theorem 4.3.5.** There exists a very good filtered extension of the original space such that we have the stable convergence in law $L_{\lfloor Nt \rfloor}^N \rightarrow L_t$ with

$$dL_t = -\mu_t \Phi_{\sigma} (\lambda^{-1} \cdot) \, dt + \sqrt{\Phi_{\sigma_t} (\gamma'_t)^2 (\lambda^{-1} \cdot)} dW'_t$$

(4.27)

where

$$d[p, L]_t = -\Phi_{\sigma_t} (id \gamma'_t (\lambda^{-1} \cdot)) \, dt.$$  

(4.28)

### 4.3.3 A special case

A flat order book corresponds to $\gamma''_t = m_t$ for some adapted process $m$. While quite unrealistic, it is *extremely* tractable and has been proposed and used in other models ([3, 41]).
This corresponds to quadratic transaction costs and linear price impact:

\[
\begin{align*}
    dp_t &= -\frac{\lambda}{m_t} dL_t \\
    dX_t &= L_t dp_t + \left(\frac{1}{2} - \lambda\right) \frac{L_t^2}{m_t} dt
\end{align*}
\]  \tag{4.29}

Note that the sign of the effective transaction costs is that of \( \frac{1}{2} - \lambda \). Indeed, in the self-financing case \( \lambda = \frac{1}{2} \), price recovery and price impact perfectly cancel each other out. If \( \lambda > \frac{1}{2} \), then the price impact of trades is stronger than the collected spread because of insufficient price recovery. Also, because of the uniform structure of the order book and perfect fill rate, the inventory of the provider is perfectly anti-correlated to the price.
Chapter 5

Applications

Having covered both the theory and the empirical analysis of the self-financing wealth equation, we turn our attention to its applications. This is a crucial point, as one of the main features of the frictionless wealth equation is its tractability, which is then carried over to some of its applications. The goal in this chapter is to show that our more accurate self-financing wealth equation remains tractable, as well as opens up problems that could not be tackled in a frictionless world.

In section 5.1 we illustrate tractability of our framework by revisiting the standard Black and Scholes option pricing problem in a local volatility world. Beyond providing more accurate pricing that incorporates both transaction cost and price impact, our result also provides a novel insight on when to delta-hedge via limit orders or market orders. We deem our pricing formulas to be equally tractable to the standard frictionless case, given that the frictionless formulas still hold with a slightly modified volatility term. The derivation method is nearly identical to the standard frictionless case and does not involve additional technicalities.

Section 5.2 covers a problem dear to the microstructure literature: absence of price manipulation strategies in a price impact model. This is a problem that can only be
asked in a setting with trading frictions and which can quickly become intractable. We pick a price impact model for which the result is easiest to prove.

Market making, covered in section 5.3, was historically the primary motivation of the author for studying the self-financing equation. It is a problem that is very relevant in high frequency finance, both from a practical trading and from a regulatory perspective. Market frictions are at the heart of the problem, making it rather dependent on the model used for the microstructure of the market. We therefore build one up from scratch using the same micro-to-macro approach preached throughout the dissertation in order to tailor it as closely as possible to our exact problem.

5.1 Option pricing

In this section we want to derive a pricing PDE for a European option in a local volatility model with transaction costs and price impact.

5.1.1 Mathematical setup

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space and $W$ be a Brownian motion which generates the filtration $\mathbb{F}$. Assume the midprice $p_t$ to satisfy the stochastic differential equation (SDE)

$$dp_t = \mu(p_t)dt + \sigma(p_t)dW_t,$$

(5.1)

and assume the order book shape function to be random only through the price level $p$: $\gamma_\alpha(\alpha) = \gamma(p_t, \alpha)$. The functions $\mu$, $\sigma$ and $\gamma$ are assumed to be continuous. We single out a trader. Assume that her inventory can attain any $\mathcal{F}_t$-adapted Itô process

$$L_t = L_0 + \int_0^t b_u du + \int_0^t l_u dW_u$$

(5.2)
where \( l_t \) is signed. We impose the reduced form trading constraint \( l_t < 0 \) when trading with limit orders. This implies that when \( l_t \geq 0 \) the inventory is obtained by trading with market orders.

Denote by \( c(p, \cdot) \) the Legendre transform of \( \gamma(p, \cdot) \). We will use the function

\[
g(p, l) = \text{sign}(l) \Phi_t(c(p, \cdot))
\]  

(5.3)

Given a real \( K_0 \) representing the trader’s initial cash endowment, her wealth is

\[
X_t = L_0 p_0 + K_0 + \int_0^t L_u dp_u + \int_0^t (\sigma(p_u) l_u - g(p_u, l_u)) du + \int_0^t r(X_t - p_t L_t) dt
\]

(5.4)

where \( r \) is the continuously compounded interest rate. This follows from our self-financing equation and the definition of continuously compounded interest on the cash account \( K_t = X_t - p_t L_t \).

Let \( f \in C^0 \) be the payoff function of a European option with maturity \( T \). We define a perfect replication strategy as follow

**Definition 5.1.1.** An initial cash endowment \( K_0 \) and an inventory process \( L_t \) are said to perfectly replicate the European payoff \( f(p_T) \) at maturity \( T \) if

\[
X_T = f(p_T)
\]

(5.5)

and the replication price is defined as \( X_0 = K_0 + p_0 L_0 \).

**Remark 5.1.2.** Note that the distribution of the initial endowment in terms of asset position and cash matters. This is due to market frictions. When quoting a price for the option, the trader therefore also needs to quote an initial delta asked of the buyer of the option.
5.1.2 The result

**Theorem 5.1.3.** Let \( f \in C^0 \) and \( T > 0 \). Assume \( v \in C^{1,2} \) to be the solution to the PDE

\[
\frac{\partial v}{\partial t}(t,p) + g \left( p, \sigma(p) \frac{\partial^2 v}{\partial p^2}(t,p) \right) - \frac{\sigma^2(p)}{2} \frac{\partial^2 v}{\partial p^2}(t,p) + rp \frac{\partial v}{\partial p}(t,p) = rv(t,p) \tag{5.6}
\]

with terminal condition \( v(T,p) = f(p) \). Then \( L_t = \frac{\partial v}{\partial p}(t,p_t) \), and the choice \( K_0 = v(0,p_0) - \frac{\partial v}{\partial p}(0,p_0)p_0 \) leads to a perfect replication strategy for the payoff \( f(p_T) \) at maturity \( T \). Furthermore,

\[
l_t = \sigma(p_t) \frac{\partial^2 v}{\partial p^2}(t,p_t). \tag{5.7}
\]

and the replication price of the option is

\[
X_0 = v(0,p_0) \tag{5.8}
\]

**Proof.** For a given trading strategy, define the discounted wealth process \( \tilde{X}_t = e^{-r(t-T)}X_t \). Choosing \( L_t = \frac{\partial v}{\partial p}(t,p_t) \) leads to

\[
l_t = \sigma(p_t) \frac{\partial^2 v}{\partial p^2}(t,p_t) \tag{5.9}
\]
and hence

\[ d\tilde{X}_t = e^{-r(t-T)} ((\mu(p_t) - rp_t) L_t + \sigma(p_t) l_t - g(p_t, l_t)) dt + e^{-r(t-T)} \sigma(p_t) L_t dW_t \]

\[ = e^{-r(t-T)} \left( (\mu(p_t) - rp_t) \frac{\partial v}{\partial p}(t, p_t) - g(p_t, \sigma(p_t) \frac{\partial^2 v}{\partial p^2}(t, p_t)) + \sigma^2(p_t) \frac{\partial^2 v}{\partial p^2}(t, p_t) \right) dt \]

\[ + e^{-r(t-T)} \sigma(p_t) \frac{\partial v}{\partial p}(t, p_t) dW_t \]

\[ = e^{-r(t-T)} \left( -r v(t, p_t) + \mu(p_t) \frac{\partial v}{\partial p}(t, p_t) + \frac{1}{2} \sigma^2(p_t) \frac{\partial^2 v}{\partial p^2}(t, p_t) + \frac{\partial v}{\partial t}(t, p_t) \right) dt \]

\[ + e^{-r(t-T)} \frac{\partial v}{\partial p}(t, p_t) \sigma(p_t) dW_t \]

\[ = d(e^{-r(t-T)} v(t, p_t)) \]

As the initial values match, we have that \( X_t = v(t, p_t) \). This concludes.

We recall that an option is said to have positive gamma when \( \frac{\partial^2 v}{\partial p^2}(t, p) > 0 \) for all \( t \) and \( p \). A negative gamma option is one for which \( \frac{\partial^2 v}{\partial p^2}(t, p) < 0 \) for all \( t \) and \( p \).

**Corollary 5.1.4.** Positive gamma options can only be hedged with market orders. Negative gamma options can only be hedged with limit orders.

**Proof.** The identity

\[ l_t = \sigma(p_t) \frac{\partial^2 v}{\partial p^2}(t, p_t) \quad (5.10) \]

implies that \( l_t \) and the option gamma must be of the same sign.

**Remark 5.1.5.** For an intuitive example of this phenomenon consider a call option. It has positive gamma: you need to buy when the price goes up and sell when the price goes down to rebalance your delta-hedge. But, because of adverse selection, the price tends to move down when you buy with limit orders and up when you sell with limit orders. Therefore, market orders are to be preferred to delta-hedge a call option. However, if you are selling a call option then the opposite of a call option must be delta-hedged, which can be done with limit orders.
The general PDE is non-linear in the second derivative of $v$. In the special case where all the trades happen at the best bid or ask price, we recover a linear PDE.

**Corollary 5.1.6.** Assume all trades to happen at the best bid or ask price. Denote by $s_t = s(p_t)$ the associated bid-ask spread. Then we have that the pricing PDE becomes

$$\frac{\partial v}{\partial t}(t,p) + \left(\frac{\sigma(p)s(p)}{\sqrt{2\pi}} - \frac{\sigma^2(p)}{2}\right) \frac{\partial^2 v}{\partial p^2}(t,p) + rp\frac{\partial v}{\partial p}(t,p) = rv(t,p) \quad (5.11)$$

**Proof.** In the bid-ask spread case, $c(p,l) = s(p)^2|l|$ and hence $g(p,l) = \frac{s(p)}{\sqrt{2\pi}}l$. \qed

In particular, we recover the standard, frictionless local volatility model when $s(p) = \sqrt{2\pi}\sigma(p)$. This is because transaction costs and price impact exactly cancel out in this case, and the frictionless self-financing equation holds true.

In addition to computing the price and delta-hedging ratios under transaction costs and instantaneous adverse selection, this theory suggests an execution strategy by quantifying when limit or market orders should be used to hedge an option.

### 5.2 Absence of price manipulation

A major concern for any dynamic model of market microstructure are price manipulation strategies. These can be defined in multiple ways and have been studied extensively by Schied et al in [4]. We will define price manipulation in the following way:

**Definition 5.2.1 (Price manipulation).** Let $c$ be a transaction cost function, $\lambda$ a price recovery parameter as in section 4.3.2 and $\mathcal{A}$ a set of admissible $(p,L)$ where the inventory is attained via limit orders. We say that this set $\mathcal{A}$ is subject to price manipulation if there exists a $(p,L)$ such that $L_1 = L_0$ and

$$E[X_1] < E[X_0] \quad (5.12)$$
In words, we want to exclude round-trip statistical arbitrages, given that the liquidity provider does not control the incoming market orders.

The tractability of the flat order book allows us to rule out such price manipulation strategies.

**Proposition 5.2.2.** Consider a market with deterministic price recovery parameter \( \lambda \leq 1 \). A constant flat order book, \( \gamma_t'' = c \) for some \( c > 0 \) does not allow for price manipulation.

**Proof.** Recall that the flat order book with deterministic price recovery corresponds to quadratic transaction costs and linear price impact:

\[
\begin{align*}
  dp_t &= -\frac{\lambda}{c} dL_t \\
  dX_t &= L_t dp_t + \left(\frac{1}{2} - \lambda\right) \frac{L_t^2}{c} dt
\end{align*}
\]  

(5.13)

The case \( \lambda = 1 \), absence of price recovery, dominates all the other cases and we can without loss of generality assume \( \lambda = 1 \). In that case, we have that

\[
dX_t = L_t dp_t - \frac{L_t^2}{2c} dt = -\frac{1}{2c} d \left( L_t^2 \right)_t
\]

and hence

\[
X_t = X_0 - \frac{1}{c} \left( L_t^2 - L_0^2 \right)
\]

(5.14)

which concludes.

\[\square\]

### 5.3 Market making

In this section, we adapt to our framework the key insight of the model proposed in \[^9\]. The ultimate aim is to solve the optimization problem of a representative market
maker choosing the spread and maximizing his profits. The trade-off he faces, and which is the key ingredient of the model, is the following: the smaller the spread, the likelier trades are, but the less profit he makes on each of them.

In a way similar to [9, 47], we model the probability of execution of a limit order by a decreasing function of the quoted spread. This will first be done at the microscopic level, to obtain a reasonable model for our inventory process $L$ at the macroscopic level. A key difference with [9] is that we still impose the price impact constraint, which will further depress the market maker’s profits because of adverse selection.

5.3.1 Microscopic model

To guarantee the price impact constraint is satisfied, we use, at the microscopic level, a modified version of the Almgren and Chriss model [6] to relate the price to the aggregate inventory of the liquidity providers. We assume that

$$
\Delta_n L = -\lambda_{n+1} \Delta_n p
$$

(5.15)

for a $\mathcal{F}_{n+1}$-measurable, positive random variable $\lambda_{n+1}$. This is an unpredictable form of linear price impact, in the sense that, ex-post, the price increment is a linear function of the traded volume.

To capture the insight of [9], we model $\lambda_{n+1}$ in such a way that

$$
\mathbb{E}[\lambda_{n+1} | \mathcal{F}_n] = \rho_n(s_n) f_n(s_n); \quad \mathbb{E}[\lambda_{n+1}^2 | \mathcal{F}_n] = (f_n(s_n))^2
$$

(5.16)

where $s_n$ is the market maker’s chosen spread, and $\rho_n$ and $f_n$ are continuous, positive function with $f_n$ decreasing and $\rho_n \in [0, 1]$. The assumption that $f_n$ is decreasing in the spread is inherited from [9], and the fact that $\rho$ must be smaller than 1 is due to Jensen’s convexity inequality. We assume $\lambda_{n+1}$ to be independent of $\Delta_n p$ conditional
on $\mathcal{F}_n$. Computing the predictable quadratic variation of $L_n$ yields:

$$\sum_{k=1}^{n-1} f_k^2(s_k)\mathbb{E}\left[\Delta_k p^2 \mid \mathcal{F}_k\right],$$

while the predictable quadratic covariation of $L_n$ and $p_n$ is given by:

$$-\sum_{k=1}^{n-1} \rho_k(s_k)f_k(s_k)\mathbb{E}\left[\Delta_k p^2 \mid \mathcal{F}_k\right].$$

### 5.3.2 Control problem on the macroscopic limit

This suggests the use of the following model in the continuum limit:

$$\begin{cases}
dp_t = \mu_t dt + \sigma_tdW_t \\
dL_t = -\rho_t(s_t)f_t(s_t)\mu_t dt + f_t(s_t)\sigma_t dW^\perp_t
\end{cases}$$

with $d[W, W^\perp]_t = -\int_0^t \rho_u(s_u) du$ for some adapted, continuous and positive functions $\rho_t(\cdot)$ and $f_t(\cdot)$ with $\rho_t \leq 1$ and $f_t$ decreasing. Note that the equation for $L_t$ can also be rewritten as:

$$dL_t = -\rho_t(s_t)f_t(s_t)dp_t + f_t(s_t)\sqrt{1 - \rho_t^2(s_t)}\sigma_t dW^\perp_t$$

with a Wiener process $W^\perp_t$ independent from $W_t$. We will from now on assume that $p_t$ is adapted to the filtration generated by $W_t$.

**Remark 5.3.1.** Equation (5.20) decomposes the market maker’s inventory into two parts: 1) a flow of orders that is perfectly anti-correlated to the price 2) a noise flow. This decomposition can again be interpreted as informed and uninformed flow.
Applying our wealth equation, we obtain:

\[ X_T = L_T P_T - \int_0^T p_t dL_t + \frac{1}{\sqrt{2\pi}} \int_0^T \sigma_t s_t f_t(s_t) dt. \]  

(5.21)

For both \( f_t \) and \( \rho_t \), a natural assumption is that they are functions of the spread rescaled by the volatility:

\[ f_t(s) = f(s/\sigma_t); \quad \rho_t(s_t) = \rho(s_t/\sigma_t) \]  

(5.22)

for some \( C^0 \) decreasing function \( f \) and \( C^0 \) function \( \rho \). We will furthermore assume that \( g(x) = xf(x) \) is a decreasing function for \( x \) large enough, that \( g(x) \to 0 \) as \( x \to \infty \), and that \( f(x) > 0 \) for all \( x \geq 0 \).

The problem of a risk-neutral market maker attempting to set the spread optimally is to maximize:

\[ \sup_s \mathbb{E} X_T. \]  

(5.23)

5.3.3 Solution

We solve this control problem using the Pontryagin maximum principle (see [44]). Let us define a few functions first.

**Lemma 5.3.2.** For all \( a > 0 \), define the function \( F_a \) by

\[ F_a : x \mapsto \frac{x}{\sqrt{2\pi} f(x)} - a \rho(x) f(x) \]  

(5.24)

Then the function

\[ M(a) = \max_{x \in [0, \infty)} F_a(x) \]  

(5.25)
is well defined, continuous, and decreasing in \( a \). Furthermore, there exist a measurable selection
\[
m(a) \in \arg \max_{x \in [0, \infty)} F_a(x)
\] (5.26)
and we have that \( m(a) > 0 \).

**Proof.** First, note that for all \( a > 0 \),
\[
F_a(0) = -a \rho(0) f(0) \leq 0, \quad F_a(a + 1) \geq f(a + 1) > 0
\]
Next, if \( g \) is decreasing on the interval \([x_0, \infty)\), then we can define the function \( \beta(a) \) as \( g^{-1} \circ f(a + 1) \) if \( f(a + 1) \) is in \( g[x_0, \infty) \), and \( x_0 \) otherwise. \( \beta(a) \) is continuous and verifies \( f(a + 1) \geq g(x) \) for all \( x \in (\beta(a), \infty) \).

This proves that the maximum of \( F_a \) is attained on the compact \([a + 1, \beta(a)]\). The continuity of \( M \) holds by Berge’s maximum theorem. It is decreasing by definition of \( F_a \). The measurable selection result follows by Thm 18.19 of [5].

**Proposition 5.3.3.** Any solution of the control problem is of the form
\[
\frac{s_t}{\sigma_t} = m(\alpha_t)
\] (5.27)
where
\[
\alpha_t = \mathbb{E} [p_T - p_t | \mathcal{F}_t] \frac{\mu_t}{\sigma_t^2} + \frac{Z_t}{\sigma_t},
\] (5.28)
\( Z_t \) being the volatility of the martingale representation of \( p_T \)

**Proof.** We apply the necessary part of the stochastic Pontryagin maximum principle. The generalized Hamiltonian is equal to:
\[
\mathcal{H}_t(s, L, Y, Z, Z) = -\rho(s/\sigma_t)f(s/\sigma_t)[(Y_t - p_t) \mu_t + \sigma_t Z]
+ \frac{\sigma_t}{\sqrt{2\pi}} s f(s/\sigma_t) + \sigma_t f(s/\sigma_t) \sqrt{1 - \rho^2(s/\sigma_t)} Z
\]
and the adjoint equation is solved by

\[ Y_t = \mathbb{E} [p_T | \mathcal{F}_t] \]  \hspace{1cm} (5.29)

which, in particular, implies \( Z_t^\perp = 0 \). \( Z_t \) can be computed via the martingale representation theorem on the Brownian filtration generated by \( W_t \).

The Hamiltonian to maximize therefore becomes

\[ \sigma_t^2 F_{\alpha_t} \left( \frac{s}{\sigma_t} \right) \]  \hspace{1cm} (5.30)

and the previous lemma concludes. \( \Box \)

Beyond the optimal control, one might be interested in the dependence in \( \sigma_t \) and \( \alpha_t \) of the market maker’s expected profits as well as the volatility of his inventory. Note that a low volatility of the inventory means that the market maker has essentially pulled out of the market.

**Corollary 5.3.4.** The market maker’s expected profits and losses are

\[ \mathbb{E} \left[ \int_0^T M (\alpha_t) \sigma_t^2 dt \right] \]  \hspace{1cm} (5.31)

while the volatility of his inventory is

\[ \sigma_t f(m(\alpha_t)). \]  \hspace{1cm} (5.32)

*Proof.* The expected profits can be computed by integrating the Hamiltonian along the optimal path. The rest follows from the previous proposition. \( \Box \)

A consequence of the corollary is that the market maker is on average short \( \alpha_t \) and, for \( \alpha_t \) being fixed, long volatility.
There are now two distinct problems if one looks for tractable formulas. First, an explicit model for $p_T$ must be given for which the martingale representation term $Z_t$ can be computed. Second, one has to propose a function $g$ for which the maximal argument $m$ of $F$ can easily be characterized as a function of $\alpha_t$.

**The martingale case**

Note that the latter problem is solved when $p_t$ is assumed to be a martingale. Indeed, if we have

$$dp_t = \sigma_t dW_t$$

(5.33)

for some adapted, continuous and positive process $\sigma_t$. Then $\alpha_t = 1$ and we simply have

$$s_t = m(1)\sigma_t$$

(5.34)

circumventing the need for explicit functions $\rho$ and $f$. This result provides a theoretical argument for the empirical claim made in [49] that the spread is a linear function of volatility.

Plugging this optimal spread back into the objective function, the market maker’s expected profits and losses (P&L) are

$$M(1)\mathbb{E}\left[\int_0^T \sigma_t^2 dt\right]$$

(5.35)

In the martingale case, the market maker is therefore on average, Delta neutral, has negative Gamma but positive Vega.

**Explicit cases**

Other cases where $\alpha_t$ can be computed explicitly are:
• the Black-Scholes model

\[ dp_t = \mu p_t dt + \sigma p_t dW_t \]  \hspace{1cm} (5.36)

in which case we obtain:

\[ \mathbb{E} [p_T | \mathcal{F}_t] = p_t e^{\mu (T-t)}; \quad Z_t = \sigma p_t e^{\mu (T-t)}, \]  \hspace{1cm} (5.37)

and hence

\[ \alpha_t = \frac{\mu}{\sigma^2} (e^{\mu (T-t)} - 1) + e^{\mu (T-t)}. \]  \hspace{1cm} (5.38)

• the case of a mean reverting (Ornstein-Uhlenbock) price process

\[ dp_t = \rho (p_0 - p_t) dt + \sigma dW_t \]  \hspace{1cm} (5.39)

in which case:

\[ \mathbb{E} [p_T | \mathcal{F}_t] = p_0 + e^{-\rho (T-t)} (p_t - p_0); \quad Z_t = \sigma e^{-\rho (T-t)}, \]  \hspace{1cm} (5.40)

and hence

\[ \alpha_t = -\frac{\rho}{\sigma^2} (p_t - p_0)^2 \left( e^{-\rho (T-t)} - 1 \right) + e^{-\rho (T-t)}. \]  \hspace{1cm} (5.41)

Unlike in the martingale case, it is hard to obtain any tractable formulas without specifying a functional form for \( \rho \) and \( f \). In the case where \( \rho(x) = 1/(1 + x) \) and \( f(x) = 1/(1 + x)^2 \), the optimal spread becomes

\[ s_t = \sigma_t \sqrt{1 + 3 \alpha_t} \]  \hspace{1cm} (5.42)
Note that $m$ is an increasing function of $\alpha_t$. To compare with the martingale case, where $\alpha_t = 1$, we therefore want to compare the ratio $\alpha_t$ to 1 to study the impact of the model assumptions on the market maker’s profits and inventory volatility.

- For the Black-Scholes model, $\alpha_t$ is larger than 1 for $\mu > 0$. For $\mu < 0$, there exists a critical value depending on $T$ and $\sigma$ for which this ratio flips sign.

- In the case of an Ornstein-Uhlenbock process, $\alpha_t$ is smaller than 1 iff

\[
(p_t - p_0)^2 < \frac{\sigma^2}{\rho} \tag{5.43}
\]

that is, if the current price $p_t$ isn’t too far from the long-term average $p_0$.

In line with intuition, the market maker quotes larger spreads, expects less profit, and captures less volume in the ‘momentum’ Black-Scholes model, as compared to the martingale case. In a mean-reverting market, unless the price is significantly away from its long-term trend, the market maker quotes smaller spreads, expects more profit and captures more volume than in the two other market models.
Chapter 6

A macroscopic model for market making in foreign exchange markets

The last chapter presents another model for market making. It is more elaborate than the one presented in section 5.3 and showcases specific microstructure features not present in the rest of the thesis. The main reason for the addition of these features -as well as the removal of others- is the desire to adapt the modeling procedure to a very particular problem: market making in Foreign Exchange.

The chapter begins by explaining in section 6.1 the context of the model and deriving the self-financing equation to be used. The following three sections outline a model for the limit order book based on a market maker’s control problem. The endeavor is complex enough to warrant two preparatory steps. Section 6.2 models the behavior of the market maker’s clients, and in particular how to imply a market view from their trades based on past behavior. Section 6.3 proposes a series of approximations to make the market maker’s problem tractable. Finally, section 6.4 solves the final market making control problem.
6.1 Context and self financing equation

We summarize in this section the main differences between the market model in this chapter and the rest of the thesis. One difference is the market modeled -Foreign Exchange- which leads to a qualitatively different microstructure much more heavily focused on the identity of individual traders. The other difference is a modeling choice: we decide in this section to use a more standard self-financing portfolio equation, in line with the Almgren and Chriss model and justify it through our usual limiting argument.

6.1.1 Market making in Foreign Exchange

The microstructure of the foreign exchange market shares some similarities, but also exhibits some crucial differences with the typical microstructure of US stock markets.

While centralized exchanges with limit order books akin to those of NASDAQ or the NYSE exist, most of the traded volume actually takes place in bilateral agreements between specialized market makers and their clients. These work in the following manner:

1. The market maker keeps a list of clients. He continuously quotes a limit order book to each of them. This limit order book may be different for each client.

2. Each client aggregates the limit order book quoted by multiple market makers, as well as the centralized exchange and trade at the best price available.

3. Ex-post, the market maker computes statistics on each client and updates his limit order book accordingly.

The main difference with trading on US stock markets is therefore the lack of anonymity, as liquidity providers can identify trades by different individuals, and respond accordingly. Nevertheless, transaction costs work in similar ways because of
the rules of the limit order book. This will be a crucial ingredient in our modeling, as implying private information becomes a crucial component of the market making strategy.

6.1.2 A different diffusion limit

In chapter 4, we established equation (4.9) as one possible self-financing equation consistent with a certain scaling of the order book. This scaling was chosen such that inventories had infinite variations and justified by the study of the quadratic covariation between price and inventory in chapter 3. Because of the perfect match with the data, we believe the self-financing equation (4.9) to be the correct one to use in high frequency markets.

Nevertheless, the standard assumption in the current mathematical financing literature is to assume that, in the presence of transaction costs, inventories must have finite variations. In particular, the standard Almgren and Chriss model assumes inventories to be differentiable. While inconsistent with the data, such an assumption can lead to tractable models for certain problems and fits very well with a more qualitative, approximate description of the market.

We therefore proceed to recall and justify such a self-financing portfolio equation for a scaling of the limit order book where inventories are differentiable.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space supporting a Wiener process $W$. We consider a fixed time interval $[0,1]$ and give ourselves the following $\mathcal{F}$-adapted processes for the price and inventory of a provider:

\[
\begin{align*}
  p_t & = p_0 + \int_0^t \mu_udu + \int_0^t \sigma_udW_u \\
  L_t & = L_0 + \int_0^t l_udu
\end{align*}
\]  

(6.1)
where \( p_0 \) and \( L_0 \) are \( \mathcal{F}_0 \)-measurable elements of \( L^2 \) and \( \mu, \sigma \) and \( l \) are \( \mathcal{F} \)-adapted and càdlàg processes. Finally, let \( c : \Omega \times [0, 1] \times \mathbb{R} \to \mathbb{R}^d \) be a random, \( \mathcal{F}_t \)-adapted function that is \( C^0 \) in both its variables. Assume \( c \) to be a.s. convex for all \( t \), with a minimum at \( c_t(0) = 0 \) and such that \( c_t(l) < Cl^2 \) for some constant \( C \). We denote by \( \gamma_t \) its Legendre transform, which will represent the shape function of the order book as measured in tick size.

Let \( \frac{1}{\sqrt{N}} \) be a vanishing tick size. Define the discretized price process as \( p_n^N = \frac{p_n}{N} \) and likewise \( L^N \). Note that in this setting, \( \Delta_n L^N \) is of order \( 1/N \) instead of \( 1/\sqrt{N} \).

In this context the proper renormalization for the transaction costs is

\[
c_n^N(l) = \frac{1}{N} c_{n/N} (Nl)
\]

which leads to the discretized self-financing equation

\[
\Delta_n X^N = L^N_n \Delta_n L^N + \frac{1}{N} c_{n/N} (N\Delta_n L^N) + \Delta_n p^N \Delta_n L^N
\]

for trading with market orders.

**Theorem 6.1.1.** The continuous time relationship between provider wealth \( X \), inventory \( L \), price \( p \) and transaction costs \( c \) is:

\[
dX_t = L_t dp_t - c_t(l_t) dt
\]

where \( X_t = \lim_{N \to \infty} X^N_{\lfloor Nt \rfloor} \) u.c.p.

**Proof.** The first term of the discrete equation converges as an Itô integral. The second term converges as a standard Riemann integral while the last term converges to zero as the empirical quadratic covariation between \( L \) and \( p \). \( \Box \)
6.2 Client control problem

In this section we define the overall setting as well as the working model for client behavior. This model is based on a simple control problem for a client. It is used by the market maker in section 6.3 to simulate future client behavior and their relationship to the midprice.

6.2.1 Heterogeneous beliefs

Consider \( n \) clients and one market maker. We will denote by \( i \in \{1...n\} \) a client index, and \( k \in \{0...n\} \) a generic index, with \( k = 0 \) corresponding to the market maker. We first introduce the following setting for the model:

1. a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\) representing the “real life” filtration and probability measure. The filtration is generated by a \( d \)-dimensional \( \mathbb{P} \) Wiener process \( W_t \).

2. a different filtration and measure \(( (\mathcal{F}_t^k)_{t \geq 0}, \mathbb{P}^k) \) for each of the agents. Assume furthermore that \( \mathcal{F}_t^k \subset \mathcal{F}_t \), that \( \mathbb{P}^k |_{\mathcal{F}_t^k} \) and \( \mathbb{P} |_{\mathcal{F}_t^k} \) are equivalent and that \( \mathbb{P}^0 |_{\mathcal{F}_t^0} = \mathbb{P} |_{\mathcal{F}_t^0} \).

3. a \( d^k \)-dimensional \( \mathbb{P} \) Wiener process \( W_t^k \) that generates the filtration \((\mathcal{F}_t^k)_{t \geq 0}\).

4. a price process \( p_t \) which is an Itô process adapted to all the filtrations \( \left( (\mathcal{F}_t^k)_{t \geq 0} \right)_{k = 0...n} \).

5. the drift and volatility of \( p_t \) grow at most polynomially in \( t \) under all probability measures.

where the last hypothesis must be understood in the a.s. and \( L^2 \) sense.

Let

\[
    dp_t = a_t dt + \sigma_t dW_t
\]

(6.5)
be the Itô decomposition of $p_t$ under $\left( (\mathcal{F}_t)_{t \geq 0}, \mathbb{P} \right)$. Hypotheses 2 and 3 imply that there exists an $(\mathcal{F}_t^k)_{t \geq 0}$ adapted process $r_t^k$ such that

$$W_t^k = B_t^k + \int_0^t r_s^k \, ds$$

(6.6)

for some $\mathbb{P}^k|_{\mathcal{F}_t}$ Wiener process $B_t^k$ and with $r_t^0 = 0$. Because $W^k$ is $d_k$ dimensional, so are $r_t^k$ and $B_t^k$. Furthermore, by the martingale representation theorem, given that $W_t^k$ is a $\mathbb{P}$ martingale, there exists an $(\mathcal{F}_t)_{t \geq 0}$ adapted, $d_k \times d$ dimensional matrix $\Sigma_t^k$ such that

$$dW_t^k = \Sigma_t^k dW_t$$

(6.7)

Finally, agent $k$ has the following view on the market under $\left( (\mathcal{F}_t^k)_{t \geq 0}, \mathbb{P}^k \right)$:

$$dp_t = a_t^k \, dt + \sigma_t^k dB_t^k$$

(6.8)

with $\sigma_t = \sigma_t^k \Sigma_t^k$ and $a_t^k = a_t + \sigma_t^k r_t^k$. In particular, $\sigma_t$ needs to live in the intersection of the images of all the $(\Sigma_t^k)^T$ for $p_t$ to be adapted to all the filtrations. Note that the $d_k$ is allowed to differ from one agent to another, in which case the $\sigma_t^k$ must be of different dimensions and hence differ.

In conclusion, all the agents have views on the price process that potentially conflict with each other’s probability measure and even filtration, but can be compared coherently within the larger probability space $\left( \Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P} \right)$.

**Example 6.2.1.** Consider the case where you have three Wiener processes $W^1$, $W^2$ and $W^3$. Let $p_t = W_t^1 + W_t^2 + W_t^3$, $\mathcal{F}_t^i = \sigma \left( (W_s^i)_{s \leq t}, (p_s)_{s \leq t} \right)$ and $r_t^i = f(W_t^i)$. You
then have that:

\[ dp_t = dW^1_t + dW^2_t + dW^3_t \text{ under } \mathbb{P} \]

\[ = f(W^1_t)dt + dB^{1,1}_t + dB^{1,2}_t \text{ under } \mathbb{P}^1 \]

\[ = f(W^2_t)dt + dB^{2,1}_t + dB^{2,2}_t \text{ under } \mathbb{P}^2 \]

where the two last decompositions are not adapted with respect to the other filtration, but the price process is nevertheless adapted to both filtrations.

6.2.2 Trade assumptions

We now allow trades among agents. All clients have a cumulative position \( L^i \) in the asset, which starts off at 0 at the beginning of the trading period. For the market maker, we rescale all the quantities according to the number of clients, hence \( L^0_t \) is his average cumulative position per client. Clients control their position through its first derivative, \( l^i_t \) but incur transaction costs. The market maker, on the other hand, has no direct control over his position, but receives the liquidity fee \( c_t \). To be precise, the market maker announces to his clients a transaction cost function \( l \mapsto c_t(l) \), which denotes the price of trading at speed \( l \) at moment \( t \). A client then chooses her preferred trading volume \( l^i_t \), and pays \( c_t(l^i_t) \) in total transaction fees. We make the following hypotheses on these two processes:

1. The trade volume process \( l^i_t \) for each client is adapted to \( (\mathcal{F}^i_t)_{t \geq 0} \) and \( (\mathcal{F}^0_t)_{t \geq 0} \).

2. The cost function process \( c_t \) is adapted to all filtrations.

3. Marginal costs are defined: \( c'_t \) is almost everywhere continuous.

4. Clients may choose not to trade, \( c_t(0) = 0 \) and the mid-price is well defined at \( p_t, c'_t(0) = 0 \).
5. Marginal costs increase with volume: $c_t$ is convex.

6. The order book distribution $\gamma''_t$ has total mass bounded by one.

Hypotheses 1 and 2 describe the information the different agents have access to. 3-5 are intuitive properties that the cost function must verify to make sense in terms of transaction costs. Finally, the hypothesis 6 is there to maintain a compact set for the clients’ controls.

### 6.2.3 The control problem

In this section, we solve the control problem for one generic client. Because her decisions have no impact on either the midprice $p_t$ or the transaction costs $c_t$, it will not affect any of the other clients’ decisions. Her control variable is not even adapted to their filtration. Similarly, her own control problem is not affected by the other clients’ decision. This allows us to drop the index $i$ in this subsection. Let $\left(\tilde{\mathbb{P}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}\right)$ denote that client’s probability measure and filtration and $\tilde{W}$ her Wiener process. The client tries to solve the following control problem:

- An admissible control $l_t$ is a stochastic process adapted to $(\tilde{\mathcal{F}}_t)_{t \geq 0}$ that lives in the support of $c_t$. Because of the hypothesis $\langle \gamma''_t, 1 \rangle = 1$, we have that the support of $c_t$ is included in $[-1, 1]$, which means that the control set $A$ is bounded.

- The state variable $L_t$ verifies the dynamics:

$$dL_t = l_t dt$$

(6.9)

- The objective function is

$$\sup_t \mathbb{E}_{\tilde{\mathbb{P}}} \left[ \int_0^\infty e^{-\beta t} (L_t dp_t - c_t (l_t) dt) \right]$$

(6.10)
We assume \( \beta \) large enough for the problem to be well defined.

**Remark 6.2.2.** The expectation to be maximized corresponds to the terminal wealth of the client at an exponentially distributed terminal time. To see this, consider an independent exponential random variable \( \tau \) with mean \( 1/\beta \). Note that

\[
E_{\tilde{P}}[X_{\tau}] = \int_0^\infty \beta e^{-\beta t} E[X_t] dt = E\left[\int_0^\infty e^{-\beta t} dX_t\right]
\]

by independence of \( \tau \) and integration by parts. The choice of an exponential terminal time was used to obtain a memory-less model for clients, a useful property from the market maker’s perspective.

Using the notation introduced in the previous section, we have the central result:

**Theorem 6.2.3. (Implied alpha)**

A client who trades optimally will follow the relationship

\[
\alpha_t = E_{\tilde{P}}\left[\int_t^\infty e^{-\beta (s-t)} dp_s \bigg| \tilde{F}_t\right]
\]

(6.11)

where \( \alpha_t \) is defined by the codebook \( c'_t(l_t) = \alpha_t \).

**Proof.** The aim now is to apply the Pontryagin maximum principle to the above control problem. The gain function is not in standard form, but a simple integration by parts solves this issue.

\[
\int_0^\infty e^{-\beta t} L_t dp_t = [e^{-\beta t} L_t p_t]_0^\infty - \int_0^\infty e^{-\beta t} L_t dp_t + \int_0^\infty e^{-\beta t} L_t \beta p_t dt.
\]

(6.12)

The first term is equal to 0 because we assume \( L_0 = 0 \) and \( \lim_{t \to \infty} e^{-\beta t} L_t p_t = 0 \) by the linear growth of \( L_t \) and the polynomial growth of \( p_t \). The client therefore maximizes:

\[
E_{\tilde{P}}\left[\int_0^\infty e^{-\beta t} \left( (\beta L_t - l_t) p_t - c_t(l_t) \right) dt\right]
\]

(6.13)
We write out the generalized Hamiltonian of the system:

\[ H(t, \omega, L, l, Y) = lY + e^{-\beta t} \left( (\beta L - l)p_t - c_t(l) \right) \]  \hspace{1cm} (6.14)

The generalized Hamiltonian is linear in \( L \) and concave in \( l \) by convexity of \( c, \) and therefore overall concave in \((L, l).\)

\( p_t \) and \( c_t \) are exogenously defined Itô process. Then, the backward equation

\[-dY_t = \beta p_t e^{-\beta t} dt - Z_t \cdot d\tilde{W}_t \]  \hspace{1cm} (6.15)

has a unique solution

\[ Y_t = \mathbb{E}_{\tilde{\mathbb{P}}} \left[ \int_t^\infty \beta p_s e^{-\beta s} ds \mid \mathcal{F}_t \right] \]  \hspace{1cm} (6.16)

By polynomial growth of the volatility of \( p_t, \) the \( Z \) term of the Backward Stochastic Differential Equation (BSDE from now on) satisfies the growth condition [A.6].

Therefore, the candidate optimal control is \( l_t^* \) verifying \( e^{\beta t} Y_t - p_t = c_t'(l_t^*). \)

This determines \( l_t = \gamma_t' \left( e^{\beta t} Y_t - p_t \right). \) Therefore the forward equation of \( L_t \) has a unique solution.

Finally, the quantity

\[ c_t'(l_t^*) = e^{\beta t} Y_t - p_t = \mathbb{E}_{\tilde{\mathbb{P}}} \left[ \int_t^\infty \beta p_s e^{-\beta(s-t)} ds - p_t \mid \mathcal{F}_t \right] \]  \hspace{1cm} (6.17)

\[ = \mathbb{E}_{\tilde{\mathbb{P}}} \left[ \int_t^\infty e^{-\beta(s-t)} dp_s \mid \mathcal{F}_t \right] \]  \hspace{1cm} (6.18)

can be seen as the “implied alpha” of the deal, that is, the difference between the projected value into the future and the current value of \( p. \)

While it makes intuitive sense\(^1\) this a non-trivial result. Indeed, \( c_t \) and hence \( l_t \) depend on the market maker’s decision, and yet \( \alpha_t = c_t'(l_t) \) becomes a quantity that

\(^1\)All the formula says is that marginal costs equal expected marginal gains.
is “intrinsic” to the client. It is independent of the market maker’s pricing and only depends on the client’s view on the market. It intuitively represents the client’s price estimator and summarizes her beliefs on the price dynamics. Note that the discount factor now becomes the time-scale of her prediction. We define the quantity the client tries to predict as the realized alpha over the time scale $\frac{1}{\beta}$:

$$\alpha_t^r = \int_t^\infty e^{-\beta(s-t)} dp_s.$$  

(6.19)

This coincides with what practitioners refer to as the 'alpha' of a client.

### 6.2.4 Dynamics of the implied alpha

(6.11) can be rewritten in terms of Itô dynamics:

$$d\alpha_t = \beta \alpha_t dt - dp_t + e^{\beta t} Z_t d\tilde{W}_t$$  

(6.20)

$$= \beta \alpha_t dt - dp_t + \theta_t d\tilde{W}_t$$  

(6.21)

with $\theta_t = e^{\beta t} Z_t$ therefore being the volatility of the estimation. All these equations summarize the link between optimal client volumes and price dynamics under the client’s probability measure and filtration.

Three things can be noted.

1. The drift of the implied alpha is the result of two opposing forces. On the one hand, the self-correlation term guarantees a certain coherence in the client’s decisions over the time-scale $\beta^{-1}$. On the other hand, the implied alpha, through the $-dp_t$ term, automatically takes into account the last price variation to recenter the estimation.

2. $\theta_t$ is a measure of intelligence of a client over the price process $p_t$, given that in the limit where $\theta_t = 0$, a client has a perfect view on the market. Conversely,
clients with a big \( \theta_t \) will have a higher variance on their price estimator, and can at the limit be considered as “noise” traders.

3. We can write the dynamics of \( \alpha_t \) under \( \mathbb{P} \) to obtain:

\[
d\alpha_t = \beta \alpha_t dt - dp_t + \theta_t \Sigma_t dW_t + \theta_t r_t dt
\]

which provides the link between trade and price dynamics. Note that, while under \( \tilde{\mathbb{P}} \), \( \alpha_t \) is intrinsic to the client, under \( \mathbb{P} \), \( \alpha_t \) may depend on the market maker’s decisions. In words, this means that while the market maker cannot influence the client’s decision under her own view of the market, he can affect that view itself.

**The cost of information**

Given the above result, it is clear that a market maker has a privileged position on the market: he catches a glimpse of everyone’s belief on the price. We can now give an interpretation of \( \gamma \) beyond the fact that \( \gamma'' \) represents the order book:

Under a martingale measure of \( p_t \), \( \gamma(l_t) \) represents the cost the client pays to the market maker for the liquidity \( l_t \), and this is the conservative interpretation of transaction costs. However, \( \gamma_t(\alpha_t) \) represents the cost the market maker pays *under the client’s view of the market*. Under \( \tilde{\mathbb{P}} \), it is as if the market maker pays the client for information on the price process.

This gives us a good intuition about the market maker’s strategy: he collects information from each client, pricing them according to the current beliefs and how much the new information brings to him: \( \gamma(\alpha) \) is essentially how much the market maker is willing to pay for a prediction of strength \( \alpha \), assuming that the client is correct.
6.2.5 Robustness analysis

In this section we generalize the implied alpha relationship to the utility function case. This illustrates the robustness and limits of the proposed codebook.

Recall the primary variables

\[
\begin{align*}
    dL_t &= l_t \, dt \\
    dK_t &= -(p_t l_t + c_t(l_t)) \, dt
\end{align*}
\]  

(6.23)

from which \( X_t \) can be reconstructed by the relationship \( X_t = p_t L_t + K_t \).

Consider an objective function of the form

\[
J = \mathbb{E} \left[ U(X, p) \right] = \mathbb{E} \left[ U(p_t L_t + K_t, p_t) \right]
\]  

(6.24)

with \( \tau \) a stopping time adapted to the client’s filtration and \( U \) her utility function. The special form guarantees that the client does not differentiate between wealth in cash and wealth in the asset. An agent concerned with the liquidity of the asset would not mark the asset to the mid.

**Remark 6.2.4.** One could generalize the utility framework to 'cash-sensitive' agents by considering a utility function of the form \( U(L, K, p) \). In this case, the below result would not hold.

The new Hamiltonian becomes:

\[
\mathcal{H}(t, \omega, L, K, l, Y_L, Y_K) = l Y_L - (p_t l_t + c_t(l)) Y_K
\]  

(6.25)

where the dual variables satisfy the BSDEs

\[
Y_{L,t} = \mathbb{E} \left[ \partial_X U(X, p_t) p_t | \mathcal{F}_t \right]
\]

\[
Y_{K,t} = \mathbb{E} \left[ \partial_X U(X, p_t) | \mathcal{F}_t \right]
\]
The optimal execution strategy therefore verifies:

\[ c'_t(L'_t) = \mathbb{E} \left[ \frac{Y_{K,t}^\tau p_t}{Y_{K,t}^\tau} \bigg| \mathcal{F}_t \right] - p_t \]  

(6.26)

which can be rewritten in a manner very similar to (6.11):

\[ \alpha_t = \mathbb{E}_Q [p_t \mid \mathcal{F}_t] - p_t \]  

(6.27)

where \( \frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{Y_{K,t}^\tau}{\mathbb{E}[Y_{X,t}^\tau]} \) is a legitimate change of measure if \( \partial_X U(X_t, p_t) \) is positive and integrable. In the case where \( \tau \) is exponential and independent of the price process, we simply recover (6.11) under a different probability measure. Given that we assume all the clients to have differing probability measures anyway, we can without loss of generality stick to (6.11).

The above computation is somewhat formal, but can be made rigorous by giving explicit integrability assumptions on \( \tau \) such that the growth condition (A.2) is verified. This is in particular the case when \( \tau \) is independent of \( p \) and has exponential moments, or if \( \tau \) is bounded.

6.3 Defining the market maker’s control problem

In this section, we work under the measure \( \mathbb{P} \) and often drop the superscript 0 referring to the market maker. Our goal is to provide an optimal market making strategy. Because of its complexity, the problem cannot be solved in full generality, and we propose a set of approximations which we justify on financial grounds.

The next four subsections identify a succession of simplifications guided by intuition based on the behavior of a typical market maker:

1. First, he should ’not hold a view on the market’. Mathematically speaking, this means that in his model for the price, the main explanatory variables are his
client’s beliefs. The error associated to this first simplification is proved to be small, though a function of the market maker’s control.

2. Second, he should model how his clients’ views might evolve into the future. A straightforward system of correlated Ornstein-Uhlenbeck processes is proposed to serve this purpose. This will be used to define an approximate objective function for the market maker.

3. The third approximation is made for mathematical convenience: we assume that the market maker has an infinite number of clients. This leads to the previous models becoming SPDEs.

4. Finally, a placeholder function is proposed to model the source of error identified in the first subsection.

6.3.1 Approximate price process

As the market maker should not hold a view on the market, we refrain from directly modeling \( p_t \) under the market maker’s measure and filtration. Instead, the market maker constructs his model from the client’s implied alphas. This is done in the following fashion:

Equation (6.22) can be turned around to describe price dynamics using the implied alpha of client \( i \) under \( \mathbb{P} \):

\[
dp_t = \beta^i \alpha^i_t dt - d\alpha^i_t + \theta^i_t \Sigma_t dW_t + \theta^i_t r_t^i dt
\]

(6.28)

and this equation holds true for all \( i \). Notice that there is no contradiction with the uniqueness of the Itô decomposition: these representations correspond to the same Itô process rewritten in terms of the variable \( \alpha^i_t \). Using a sequence of positive and
constant weights $\lambda^i$ averaging to 1 (i.e. such that $(1/n) \sum_{i=1}^n \lambda^i = 1$), we obtain:

$$dp_t = \frac{1}{n} \sum_{i=1}^n \lambda^i (\beta^i \alpha^i_t dt - d\alpha^i_t) + \left( \frac{1}{n} \sum_{i=1}^n \lambda^i \theta^i_t \Sigma^i_t \right) dW_t + \left( \frac{1}{n} \sum_{i=1}^n \lambda^i \theta^i_t r^i_t \right) dt \quad (6.29)$$

Again, this equation is just a reformulation of the Itô decomposition of $p_t$. The advantage of that particular representation is that the first term is adapted to $(\mathcal{F}_t^0)_{t \geq 0}$. Hence, if he also knows (or rather, chooses) the weights $(\lambda^i)_{i=1...n}$, then the market maker can follow the first term of this decomposition in real time. We introduce the special notation $p_t^\lambda$ for this term and we refer to it as his price estimator: So:

$$dp_t^\lambda = \frac{1}{n} \sum_{i=1}^n \lambda^i (\beta^i \alpha^i_t dt - d\alpha^i_t). \quad (6.30)$$

The remainder which we denote $\epsilon_t^\lambda$ includes quantities that are unknown to him since

$$d\epsilon^\lambda = \left( \frac{1}{n} \sum_{i=1}^n \lambda^i \theta^i_t \Sigma^i_t \right) dW_t + \left( \frac{1}{n} \sum_{i=1}^n \lambda^i \theta^i_t r^i_t \right) dt \quad (6.31)$$

If we replace $p$ by $p^\lambda$ in the market maker’s problem, the only difference appears in the objective function, which now contains an extra term. We refer to is as the error term:

$$\text{err} = \mathbb{E} \left[ \int_0^\infty e^{-\beta t} L_t \left( \frac{1}{n} \sum_{i=1}^n \lambda^i \theta^i_t r^i_t \right) dt \right] \quad (6.32)$$

on the market maker’s objective function. Next we define

$$(\sigma^i)^2 = \mathbb{E} \left[ \int_0^\infty e^{-\beta t} |\theta^i_t|^2 dt \right]. \quad (6.33)$$

This quantity is a measure of how ‘noisy’ client $i$ is, and can be estimated by fitting the expression for the implied alpha to historical data. To be more specific, recall
that under the client measure and filtration,

$$\alpha^i_t = \mathbb{E}_{P^i}\left[ \int_t^\infty e^{-\beta(t-s)} dp_s \mid \mathcal{F}^i_t \right]$$

(6.34)

and that $\theta^i_t$ is the error on this estimation. Hence, the smaller $\theta^i_t$, the closer the implied alpha is to the realized one, which means that the client is particularly well informed. Note that in the finance literature, clients are often called informed if their market impact function (their average alpha) is unusually large. Our notion of intelligence of the price process does not coincide with this practice, as one can have at the same time a systematically small though correct alpha. In this paper, an informed trader is a trader for whom the implied and realized alphas nearly coincide, whereas for a noise trader, relationship (6.11) has much higher variance (due, for example, to liquidity concerns, a strongly non-linear utility function, or a poor filtration). The $\beta^i$ and $\sigma^i$ can be estimated from historical data by regressing the implied alpha against the realized one, that is for example by solving the least squares regression problem

$$\inf_{\beta > 0} \frac{1}{N} \sum_{k=1}^N \left( \alpha^i_{t_k} - \int_{t_k}^\infty e^{-\beta(s-t_k)} dp_s \right)^2$$

(6.35)

where the $t_k$ are the times in the past at which client $i$ traded.

A skilled market maker can therefore construct his approximate price process by choosing a $\lambda$ which puts most of its weight on such intelligent clients and little weight on noise traders with large $\sigma^i$'s.

Using Cauchy-Schwartz’s inequality and the particular choice of weights

$$\lambda^i = \frac{n \left( \sigma^i \right)^{-2}}{\sum_j \left( \sigma^j \right)^{-2}},$$

(6.36)
and assuming $L_t$ is uniformly bounded by a constant $\bar{L}$, we obtain:

$$|\text{err}|^2 \leq \beta^{-1}(\bar{L})^2 \mathbb{E} \left[ \int_0^\infty \beta e^{-\beta t} \frac{1}{n} \sum_{i=1}^{n} \lambda^i \theta^i r^i_t \right]^2 dt$$

$$\leq \beta(\bar{L})^2 \frac{1}{n} \sum_{i=1}^{n} \int_0^\infty e^{-\beta t} \mathbb{E} |\lambda^i \theta^i|^2 \frac{1}{n} \int_0^\infty e^{-\beta t} \sum_{i=1}^{n} \mathbb{E} |r^i|^2 dt$$

$$\leq \beta I^{-1}(\bar{L})^2 \int_0^\infty e^{-\beta t} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} |r^i|^2 dt$$

where

$$I^{-1} = \min_\lambda \frac{1}{n} \sum_{i=1}^{n} (\lambda^i \sigma^i)^2 = \left( \frac{1}{n} \sum_{i=1}^{n} (\sigma^i)^{-2} \right)^{-1}. \tag{6.37}$$

$I$ can therefore be seen as a measure of the aggregate intelligence the clients have over the price process. Note that it suffices for one $\sigma^i$ to be of order $\epsilon$ for $I$ to be of order $\epsilon^{-1}$. By Girsanov’s theorem, we then have that

$$\mathbb{E} |r^i|^2 = -2 \frac{d}{dt} \mathbb{E} \log \frac{d\mathbb{P}^i}{d\mathbb{P}^i|\mathcal{F}_t}, \tag{6.38}$$

and finally,

$$|\text{err}|^2 \leq 2\beta^2 I^{-1}(\bar{L})^2 \int_0^\infty e^{-\beta t} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E} \log \frac{d\mathbb{P}^i}{d\mathbb{P}^i|\mathcal{F}_t} dt \tag{6.39}$$

Two important remarks are in order at this point. First, as long as at least one agent is well informed, $I^{-1}$ is small. Second, the dependence of $p$ and $\alpha^i$ upon the market maker’s control $\gamma''$ is hidden in the Radon Nykodym derivative $\frac{d\mathbb{P}^i}{d\mathbb{P}^i|\mathcal{F}_t}$. To understand why it is (unfortunately) reasonable to assume that this term may be strongly dependent upon $\gamma''$ is because a client can use the information on the order book $\gamma''$ publicly available as one of the sources of information he uses to form his probability measure on the price. This problem will be addressed in the last subsection.
6.3.2 Approximate objective function

In this subsection, we take a crucial methodological step. Instead of maximizing

\[ J = \mathbb{E} \left[ \int_0^\infty e^{-\beta t} \left( L_t dp_t + \frac{1}{n} \sum_i \alpha_i \gamma_i \left( \alpha_i \right) dt - \frac{1}{n} \sum_i \gamma_i \left( \alpha_i \right) dt \right) \right] \] (6.40)

we assume that the market maker maximizes the approximate objective function

\[ J^\lambda = \int_0^\infty e^{-\beta t} \frac{1}{n} \sum_{i=1}^n \mathbb{E} \left[ L_t \lambda_i \left( \beta_i - \beta \right) \alpha_i \lambda_i \left( \alpha_i \right) + \left( \alpha_i \lambda_i \left( \alpha_i \right) \right) \gamma_i \left( \alpha_i \right) \right] dt \] (6.41)

subject to a constraint of the form \(|\text{err}|^2 \leq C\) for some constant \(C > 0\). The new objective function was obtained by replacing \(dp_t\) by \(dp^\lambda_t\) and integrating by parts.

Because of our choice of the form of the approximate price (6.30), the market maker’s objective function does not depend upon \(p_t\) anymore, and he only needs to model the client belief distribution. This is consistent with the intuition that a market maker should not hold a view on the market. Rather, he should model the behavior of his clients with respect to each other, and price according to the information they provide. This is exactly the approach used in what follows.

However, we still need to propose a model for the \(\alpha_i\). For reasons of tractability we choose them as correlated Ornstein-Uhlenbeck processes:

\[ d\alpha^i_t = -\rho \alpha^i_t dt + \sigma dM_t^i + \nu dW_t \] (6.42)

with \(\rho > 0\) and the \(M^i\) Wiener processes which are independent of each other and of \(W\). \(\sqrt{\sigma^2 + \nu^2}\) is the overall volatility level of a client, and \(\nu\) the volatility that is due to some common information amongst clients (for example, the movement of the midprice). The next step of our strategy is to introduce a penalization term to account for the possible feedback effects hidden inside the error term introduced when replacing the objective function by its approximation. But first, we take the limit
\( n \to \infty \) to identify effective equations providing informative approximation to the properties of the original system comprising finitely many clients.

6.3.3 Infinitely many clients

Assume that the number of clients of the market maker is large enough to justify an approximation in the asymptotic regime \( n \) large. This will greatly improve the tractability of the model by allowing us to work in function spaces and rely on stochastic calculus tools to solve the model. The mainstay of this subsection is the Stochastic Partial Differential Equation (SPDE for short):

\[
dv_t = \left( \frac{1}{2} (\sigma^2 + \nu^2) \Delta v_t + \rho \nabla \left( \text{id} v_t \right) \right) dt - \nu \nabla v_t dW_t
\]

(6.43)

describing the dynamics of an infinite dimensional measure valued Ornstein-Uhlenbeck process. The following lemma links this macroscopic SPDE to our microscopic Orstein-Uhlenbeck model for the implied alphas.

**Proposition 6.3.1.** If \((e^i)_{i \geq 1}\) is a sequence of random variables such that \((\alpha^0_i, e^i)_{i \geq 1}\) is an iid sequence independent of \(W\) and the sequence \((M^i)_{i \geq 1}\), and such that the joint distribution, say \(m\), of all the couples \((\alpha^0_i, e^i)\) satisfies:

\[
\int \left( |\alpha|^p + |e|^2 \right) m(d\alpha, d\epsilon) < \infty,
\]

(6.44)

for all \( p > 0 \), then for each \( t \geq 0 \), the limit

\[
\nu_t = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} e^i \delta_{\alpha^i_t}
\]

(6.45)

exists almost surely in the sense of weak convergence of measures, almost surely for every \( t \geq 0 \) it holds:

\[
\int |\alpha|^p \, d\nu_t(\alpha) < \infty
\]

(6.46)
for every $p > 0$, and the measure valued process $(\nu_t)_{t \geq 0}$ is a weak solution of the SPDE (6.43) in the sense that for any twice continuously differentiable function $f$ (i.e. $f \in C^2$) such as $f$ and its two derivatives have at most polynomial growth, we have that

\[
d\langle f, v_t \rangle = \left\langle \frac{1}{2} (\sigma^2 + \nu^2) \Delta f - \rho idf, v_t \right\rangle dt + \nu \langle \nabla f, v_t \rangle dW_t \tag{6.47}
\]

Furthermore, for each $t > 0$, the measure $\nu_t$ possesses an $L^2$ density almost surely.

Proof. See appendix. \qed

The papers [32, 31] provide similar results for a more general class of microscopic models, including existence and uniqueness of the solution of the SPDE (6.43). However, given the simple form of the dynamics chosen in our particular model, we can provide an explicit form for the solution and detailed properties on the nature of its tails. Note that the correlation between client beliefs is crucial in having a fully stochastic model, given that for $\nu = 0$, $\nu_t$ only satisfies a deterministic partial differential equation.

We shall use four measure valued solutions of the above SPDE. In each case, the weights $\epsilon^i$ are explicit functions of the parameters $\sigma^i$ and $\beta^i$ introduced earlier. Because the values of these parameters appear as outcomes of statistical estimation procedures in practice, assuming that they are random and satisfy some form of ergodicity is not restrictive\footnote{The fact that we have to enlarge the Brownian filtration at time $t = 0$ to randomize the coefficients does not impact the martingale representation theorem.}. To construct these measures we assume that $(\alpha_0^i)_{i \geq 1}$ is an iid sequence of random variables whose common distribution has finite moments of all orders. The first of our four measures is obtained by choosing $\epsilon^i \equiv 1$ for all $i \geq 1$. Then:

\[
\mu_t = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{\alpha_i^t}. \tag{6.48}
\]
Next we assume that \( \{(\sigma^i)^{-2}\}_{i\geq 1} \) is a sequence of positive random variables of order 1 such that \( \{(\alpha^i, (\sigma^i)^{-2})\}_{i\geq 1} \) is an iid sequence independent of \( W \) and the sequence \( (M^i)_{i\geq 1} \), so by choosing \( \epsilon^i = (\sigma^i)^{-2} \) for all \( i \geq 1 \) we can define:

\[
I_t = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} (\sigma^i)^{-2} \delta_{\alpha^i}.
\]  

(6.49)

We shall also assume that the number \( I = \mathbb{E}[(\sigma^i)^{-2}] \) is finite and strictly positive, and for each \( t \geq 0 \) we define the non-negative measure \( \lambda_t \) by:

\[
\lambda_t = I_0^{-1} I_t.
\]  

(6.50)

Finally, we assume that \( (\beta^i)_{i\geq 1} \) is a sequence of bounded random variables such that \( \{(\alpha^i, (\sigma^i)^{-2}, \beta^i)\}_{i\geq 1} \) is an iid sequence independent of \( W \) and the sequence \( (M^i)_{i\geq 1} \), so by choosing \( \epsilon^i = \beta^i (\sigma^i)^{-2} \) for all \( i \geq 1 \) we can define:

\[
\beta_t = I^{-1} \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \beta^i (\sigma^i)^{-2} \delta_{\alpha^i}.
\]  

(6.51)

We now make a few remarks on the properties shared by essentially all the solutions \( (v_t)_{t \geq 0} \) of the SPDE (6.43). For the sake of definiteness, we shall assume that \( (v_t)_{t \geq 0} \) is a non-negative measure valued process solving this SPDE.

1. The total mass \( \langle 1, v_t \rangle \) of the measure \( v_t \) is constant over time. Indeed, using the constant test function \( f \equiv 1 \) in (6.43) we see that

\[
d \langle 1, v_t \rangle = 0
\]  

(6.52)

In particular, the intelligence assumption \( I = \langle 1, I_0 \rangle \geq \epsilon^{-1} \) is conserved over time.

2. If we use the test function \( f = id \) in (6.43) where the identity function \( id \) is defined by \( id(\alpha) = \alpha \), then we see that the first moment of \( v_t \) is itself an Ornstein-Uhlenbeck...
process mean reverting around 0 since:

\[ d\langle id, v_t \rangle = -\rho \langle id, v_t \rangle \, dt + \nu \langle 1, v_0 \rangle \, dW_t. \]  \quad (6.53)

(3) Using \( f = id^2 \) we see that the second moment mean reverts around \( (\sigma^2 + \nu^2) \langle 1, v_0 \rangle \) since:

\[ d\langle id^2, v_t \rangle = (\langle \sigma^2 + \nu^2, v_0 \rangle - 2\rho \langle id^2, v_t \rangle) \, dt + 2\nu \langle id, v_t \rangle \, dW_t. \]  \quad (6.54)

These SDEs guarantee the existence of a constant \( C_1 \) such that:

\[ \mathbb{E} \langle id^2, v_t \rangle \leq e^{C_1 t} \]
\[ \mathbb{E} \langle |id|, v_t \rangle \leq \sqrt{\mathbb{E} \langle id^2, v_t \rangle} \leq e^{C_1 t/2}. \]

We shall use these estimates for the measures \( \mu_t, \lambda_t \) and \( \beta_t \). In fact similar estimates hold for moments of all order as can be proved by induction from \([6.43]\). We do not give the details as they will not be used in what follows.

Coming back to the optimal control problem of the market maker, since \( \gamma''_t \) belongs to a space of probability measures whenever the control \( (\gamma_t)_{t \geq 0} \) is admissible, the following estimates hold:

\[ ||\gamma'_t||_\infty \leq 1 \]
\[ |\gamma_t(\alpha)| \leq \alpha \]
\[ |\langle \gamma'_t, v_t \rangle| \leq \langle 1, v_0 \rangle = O(1) \]
\[ \mathbb{E} |\langle \gamma_t, v_t \rangle| \leq \mathbb{E} \langle |id|, v_t \rangle = o(e^{\beta t}) \]
\[ \mathbb{E} |\langle id \, \gamma'_t, v_t \rangle| \leq \mathbb{E} \langle |id|, v_t \rangle = o(e^{\beta t}) \]

As an immediate consequence of the above remarks we have:
Corollary 6.3.2. For any progressively measurable process \((\gamma_t)_{t \geq 0}\) such that \(\gamma''_t\) is a probability measure, the state dynamic equation of the market maker:

\[
    dL_t = -\langle \gamma'_t, \mu_t \rangle dt \tag{6.55}
\]

makes sense, and for sufficiently large \(\beta\), the approximate objective function

\[
    J^\lambda = \int_0^\infty e^{-\beta t} \mathbb{E} \left[ L_t (id, \beta_t) + \langle -L_t \beta id + (id - \bar{\alpha}_t) \gamma'_t - \gamma_t, \mu_t \rangle \right] dt \tag{6.56}
\]

where \(\bar{\alpha}_t = \langle id, \lambda_t \rangle\), is well defined.

6.3.4 Modeling the error term

As argued in the first subsection, the approximation technique hinges on one hypothesis: the existence of at least one '\(\epsilon\)-intelligent' client. Furthermore, the clients probability measure (by which we mean the distribution of the clients implied alphas) is potentially a function of the market maker’s control. This causes an undesirable nonlinear feedback effect which needs to be reined in to avoid explosion of the approximation error. In this subsection we propose a direct description of the error term, leaving open the question of how to derive it from a specific model of the clients probability measures.

Intuitively, the feedback effect corresponds to how much new information the order book shape reveals to the clients. The clients’ beliefs at time \(t\) can be summarized by the probability measure \(\mu_t\) and the order book by \(\gamma''_t\). A reasonable model for the error term is given by the expression

\[
    E = \epsilon \mathbb{E} \int_0^\infty e^{-\beta t} H(\gamma''_t | \mu_t) dt \tag{6.57}
\]
where $H(\nu|\mu)$ denotes the Kullback-Leibler distance (also known as relative entropy) defined by

$$
H(\nu|\mu) = \begin{cases} 
\int f \left( \frac{d\nu}{d\mu} \right) d\mu, & \text{whenever } \nu << \mu; \\
\infty & \text{otherwise} 
\end{cases}
$$

(6.58)

with $f(x) = x \log x$. Note that $H(\nu|\mu)$ is minimal for $\gamma''_t = \mu_t$ by convexity of $f$. As explained in the introduction, we choose this particular distance for its intuitive interpretation and the fact that it leads to an explicit expression for the two hump order book shape endogenous to the model. However, the results hold for general strictly convex functions $f$, in which case the pseudo-distance $H$ defined in (6.58) is known as the $f$-divergence between the measures $\mu$ and $\nu$. See [21].

6.4 The market maker’s control problem

With all the pieces of our model in place, we now solve the market maker’s control problem.

6.4.1 Model summary

We consider a sequence $(\beta^i, (\sigma^i)^{-2})_{i \geq 1}$ of random variables such that the assumptions of Proposition 6.3.1 are satisfied with $\epsilon^i \equiv 1$, $\epsilon^i = (\sigma^i)^{-2}$ and $\epsilon^1 = \beta^i (\sigma^i)^{-2}$ respectively, so that the measure-valued processes $(\mu_t)_{t \geq 0}$, $(\lambda_t)_{t \geq 0}$ and $(\beta_t)_{t \geq 0}$ constructed in (6.48), (6.50) and (6.51) are well defined.

As explained earlier, the market maker’s control at time $t$ is the convex function $\gamma_t$, and given our choice of penalizing terms, we expect its second derivative $\gamma''_t$ to be a probability measure absolutely continuous with respect to $\mu_t$ in order to avoid infinite penalties. So we refine the definition of the set $\mathcal{A}$ of admissible controls for the market maker as the set of random fields $(g_t(x))_{t \geq 0, x \in \mathbb{R}}$ such that $(g_t(x))_{t \geq 0}$ is progressively measurable for each $x \in \mathbb{R}$ fixed, and $\langle g, \mu_t \rangle = 1$ for each $t \geq 0$. Given
an admissible control \( g \in A \), we define the function \( \gamma_t \) as the anti-derivative of the function \( \gamma'(\alpha) = \langle g_t1_{[0,\alpha]}, \mu_t \rangle \) satisfying \( \gamma_t(0) = 0 \). The objective of the market maker is to choose an admissible control in order to maximize his modified objective function defined as:

\[
\int_0^\infty e^{-\beta t} \mathbb{E} \left[ L_t \langle \text{id}, \beta_t \rangle + \langle -L_t \beta \text{id} + (\text{id} - \bar{\alpha}_t) \gamma'_t - \gamma_t - \epsilon f \circ g_t, \mu_t \rangle \right] dt \quad (6.59)
\]

where the controlled dynamics of the state \( L_t \) of the system are given by:

\[
dL_t = -\langle \gamma'_t, \mu_t \rangle dt. \quad (6.60)
\]

We used the functions \( \gamma_t \) and \( \gamma'_t \) for consistency with earlier discussions. See below for forms of the state dynamics and market maker objective function written exclusively in terms of the density \( g_t \). The above quantities are well-defined because of the estimates proven for the measures \( \mu_t, \lambda_t \) and \( \beta_t \) solving the SPDE.

### 6.4.2 Solving the market maker control problem

We denote by \( Y \) the adjoint variable of \( L \) so that the generalized (random) Hamiltonian of the problem can be defined as

\[
H(t, L, g, Y) = -\langle \gamma'_t, \mu_t \rangle Y + e^{-\beta t} \left( L \langle \text{id}, \beta_t \rangle + \langle -\beta L \text{id} + (\text{id} - \bar{\alpha}_t) \gamma'_t - \gamma_t - \epsilon f \circ g, \mu_t \rangle \right) \quad (6.61)
\]

for any deterministic function \( g \) as long as we define \( \gamma \) and \( \gamma' \) by \( \gamma'(\alpha) = \langle g_t1_{[0,\alpha]}, \mu_t \rangle \) and \( \gamma(\alpha) = \int_0^\alpha \gamma'(\tilde{\alpha})d\tilde{\alpha} \). Note that in the above expression of the Hamiltonian, \( g \) is deterministic and does not depend upon \( t \), but \( \gamma'_t \) which is the cumulative distribution function of the probability measure \( \gamma''_t \) with density \( g \) with respect to \( \mu_t \) is random.
and depends upon \( t \). Clearly, so is \( \gamma_t \). This Hamiltonian can be rewritten as

\[
H(t, L, g, Y) = e^{-\beta t} L \langle id, \beta_t - \beta \mu_t \rangle + e^{-\beta t} \tilde{H}(g)
\]  

(6.62)

where the modified Hamiltonian \( \tilde{H} \) is defined by

\[
\tilde{H}(g) = -\langle \gamma_t', \mu_t \rangle e^{\beta t} Y + \langle (id - \bar{\alpha}_t) \gamma_t' - \gamma_t - \epsilon f \circ g, \mu_t \rangle,
\]

and the other terms do not depend upon the control. Since the stochastic maximum principle proven in appendix says that we can look for the optimal control by maximizing the Hamiltonian, we shall maximize the modified Hamiltonian. For each choice of the admissible control \((g_t)_{t \geq 0}\), we consider the corresponding state process \((L_t)_{t \geq 0}\) given by (6.60) and the adjoint equation:

\[
-dY_t = e^{-\beta t} \langle id, \beta_t - \beta \mu_t \rangle dt - Z_t dW_t.
\]  

(6.63)

Since the derivative of the Hamiltonian with respect to the state variable \( L \) does not depend upon the control or the state \( L \), the adjoint process can be determined independently of the choice of the admissible control \((g_t)_{t \geq 0}\) and the associated state \((L_t)_{t \geq 0}\) solving (6.60). Given the explicit formulas we have derived for \( \mu_t \), we obtain:

**Lemma 6.4.1.** The solution to the adjoint Backward Stochastic Differential Equation (6.63) is given by:

\[
Y_t = \frac{e^{-\beta t}}{\beta + \rho} \langle id, \beta_t - \beta \mu_t \rangle,
\]  

(6.64)

and it verifies the growth condition (A.6).

**Proof.** The exact form (6.64) of the solution can be guessed by going back to the explicit system of finitely many Ornstein-Uhlenbeck processes and taking the limit. However, for the proof, we show that the process \((Y_t)_{t \geq 0}\) given by (6.64) is the solution
by direct inspection, computing the Itô differential of $Y_t$ defined by (6.64) and using the fact that the random measures $\beta_t$ and $\beta\mu_t$ also solve the SPDE (6.43). Using (6.53), we get:

$$dY_t = -\beta Y_t dt - \rho Y_t dt + e^{-\beta t} \frac{\nu}{\beta + \rho} dW_t$$

and by the growth properties of the first moment, $\lim_{t \to \infty} Y_t = 0$. Moreover, $Z_t = e^{-\beta t} \nu/((\beta + \rho)$ clearly verifies the growth condition (A.6).

The form of the modified Hamiltonian justifies the introduction of the quantity:

$$\alpha^*_t = \bar{\alpha}_t + e^{\beta t} Y_t = \langle id, \lambda_t + \frac{\beta t - \beta \mu_t}{\beta + \rho} \rangle$$  \hspace{1cm} (6.65)

so that, if we compute the modified Hamiltonian along the path of the adjoint process we get:

$$\tilde{H}(g) = -\langle \alpha^*_t \gamma'_t - \gamma_t, \mu_t \rangle - \epsilon \langle f \circ g, \mu_t \rangle.$$

$\alpha^*_t$ can be viewed as the shadow alpha of the market maker.\footnote{This terminology is justified by the fact that, when the market maker does have his own view on the market (that is, his own model for $dp_t$), we would obtain the same optimization problem replacing $\alpha^*_t$ by $E \left[ \int_t^{\infty} e^{-\beta(s-t)} dp_s \bigg| \mathcal{F}_t \right]$.}

The first term in the definition of $\alpha^*_t$ is the average belief for alpha under the weighted client measure $\lambda_t$, while the second term takes into account mismatches in the time-horizons of the clients. For each $t \geq 0$, we define the profitability function $m_t$ by:

$$\alpha \mapsto m_t(\alpha) = (\alpha - \alpha^*_t)[\mu_t([\alpha, \infty)) - 1_{(-\infty,0)}(\alpha)].$$ \hspace{1cm} (6.66)

For each $\alpha \in \mathbb{R}$, $(m_t(\alpha))_{t \geq 0}$ is a progressively measurable stochastic process and for each fixed $t \geq 0$, the function $\alpha \mapsto m_t(\alpha)$ is almost surely continuous in $\alpha$, except
for a possible jump $m_t(0^+) - m_t(0^-)$ at $\alpha = 0$. $m_t$ is bounded and vanishes at the infinitives because of the integrability property [6.46] of the solutions of the SPDE. The profitability function quantifies the expected profit for an order placed at time $t$ at the price level $\alpha$. Indeed, the absolute value $|\alpha - \alpha_*^t|$ is equal to the spread the market maker expects to gain per filled order, and up to a possible change of sign, the term $\mu_t([\alpha, \infty)) - 1_{(-\infty,0]}(\alpha)$ is equal to the proportion of clients that will fill the order. If the arrivals of the agents of our model occurred according to a Poisson process instead of simultaneously, this would be the filling probability of the order. The respective contributions of these two terms are commonly parsed by practitioners. Notice also that, in the degenerate case $\epsilon = 0$, the profitability function is the derivative of the Hamiltonian in the direction of the control.

We now identify the modified Hamiltonian in terms of the control $g$ without involving the anti-derivatives $\gamma'_t$ an $\gamma_t$.

**Lemma 6.4.2.** For each $t \geq 0$, we have the identity

$$\langle (id - \alpha_*^t) \gamma'_t - \gamma_t, \mu_t \rangle = \langle m_t, \gamma''_t \rangle. \quad (6.67)$$

**Proof.** Successive integrations by parts and simplifications yield:

$$\langle \gamma''_t, (id - \alpha_*^t) (\mu_t([\cdot, \infty)) - 1_{(-\infty,0]}(\cdot)) \rangle$$

$$= -\langle \gamma'_t, (\mu_t([\cdot, \infty)) - 1_{(-\infty,0]}(\cdot)) - (id - \alpha^*) \mu_t - \alpha_*^t \delta_0 \rangle$$

$$= \langle \gamma_t, \mu_t - \delta_0 \rangle + \langle \gamma'_t (id - \alpha^*), \mu_t \rangle + \alpha_*^t \gamma'(0)$$

$$= \langle (id - \alpha_*^t) \gamma'_t - \gamma_t, \mu_t \rangle$$
where the first integration by parts is justified by the fact that \( \gamma' \) is bounded, the Dirac distribution has compact support and the simple facts:

\[
\lim_{\alpha \to \infty} \mu_t([\alpha, \infty)) = 0, \quad \text{and} \quad \lim_{\alpha \to -\infty} \mu_t([\alpha, \infty)) - 1_{\alpha \leq 0} = -\lim_{\alpha \to -\infty} \mu_t((-\infty, \alpha]) = 0,
\]

and \( \alpha \mu(\alpha) \) vanishes at the infinities. The linear growth of \( \gamma \) justifies the second integration by parts. The last line uses the initial conditions of \( \gamma_t \) and \( \gamma'_t \).

We can now state and prove the main result of this section:

**Theorem 6.4.3.** The solution to the market maker’s control problem is given by

\[
g_t(\alpha) = \frac{e^{mt(\alpha)/\epsilon}}{\int e^{mt(\alpha)/\epsilon} d\mu_t(\alpha)}
\]

(6.68)

The quantity in the right hand side of (6.68) can be seen as the change of measure from the distribution of the clients alphas to the order book of the market maker.

**Proof.** Using Lemma 6.4.1 and equation (6.67) enables us to rewrite the modified Hamiltonian as:

\[
\tilde{H}(g) = \langle m_t g, \mu_t \rangle - \epsilon \langle f \circ g, \mu_t \rangle
\]

(6.69)

which we need to maximize for each fixed \( t \geq 0 \), over \( g \in L^1(\mu_t) \) with \( \langle g, \mu_t \rangle = 1 \) and \( g \geq 0 \). Again, for each fixed \( t \) and almost surely, by convexity of \( f \), \( \tilde{H} \) is concave in \( g \) and bounded from above, so that \( \tilde{M} = \sup_{A_t} \tilde{H}(g) \) exists and is finite.

Let \( (g_n)_{n \geq 1} \) be a maximizing sequence in \( A_t \), i.e. a sequence of admissible \( g_n \) such that \( H(g_n) \to \tilde{M} \). By Komlos’s lemma (see for example [22]), there exists a subsequence \( g_{\phi(n)} \) and an element \( g_\infty \in L^1(\mu_t) \) such that \( \frac{1}{n} \sum_{i=1}^n g_{\phi(i)} \to g_\infty \) \( \mu_t \)-a.e. as \( n \to \infty \) for any subsequence of the original subsequence. Clearly, \( g_\infty \geq 0 \) \( \mu_t \) almost everywhere. To prove \( \langle g_\infty, \mu_t \rangle = 1 \) we show uniform integrability of \( \{g_n\}_n \) using De la Vallée-Poussin’s theorem (see [40]). Indeed, by definition of the maximizing sequence,
\( H(g_n) \geq \tilde{M} - 1 \) for large enough \( n \). Hence \( \epsilon \langle f \circ g_n, \mu_t \rangle \leq -\tilde{M} + 1 + ||m_t||_{\infty} \). \( f \) is non-negative, convex, increasing on \([1, \infty)\) and verifies \( \lim_{x \to \infty} \frac{1}{x} f(x) = \infty \). This concludes the proof of the admissibility of \( g_{\infty} \) as a control. Finally, by concavity of the modified Hamiltonian, the supremum is attained at \( g_{\infty} \).

We now characterize the maximal element we just constructed. We introduce a Lagrange multiplier \( \eta \in \mathbb{R} \) to relax the problem to the set of \( g \in L^1(\mu_t) \) such that \( g \geq 0 \), so that the Lagrangian reads:

\[
\tilde{L}(g) = \langle (m_t - \eta)g - \epsilon f \circ g, \mu_t \rangle
\]  

(6.70)

Classical variational calculus yields the optimality condition

\[
g = (f')^{-1} \left( \epsilon^{-1} (m_t - \eta) \right),
\]  

(6.71)

for some \( \eta \). The existence and uniqueness of a Lagrange multiplier renormalizing \( g \) is obvious given the explicit formula for \( (f')^{-1} \). We conclude that \( g \) is the desired optimum.

\[\square\]

### 6.4.3 Interpretation

Figure 6.1 and Figure 6.2 illustrate what the market maker does:

1. He tries to stick to a shape not too far away from \( \mu_t \), his client alpha distribution. This is to avoid feedback effects and associated errors on the price estimation.

2. He also takes into account the profitability function \( m_t \), which leads him “to make a big hole” in the center of the distribution \( \mu_t \).

The combined effects lead to the familiar “double hump” shape of the order book, as seen in [49]. Other consequences of the liquidity formula are
Figure 6.1: Graph of a simulated profitability curve $m$ with $\alpha^* \approx -0.33$. The maximum profitability is only 5% of the volatility used to simulate prices. The average spread being comparable to the volatility (see [49]), this means that the market maker’s margins are very thin in this simulation.

Figure 6.2: Bimodal distribution: Optimal order book $\gamma''$ for $\epsilon = 0.01$ using the simulated alpha distribution and the entropic penalizing factor. Unimodal distribution: Client alpha distribution. The profitability used is the same as in figure 6.1.
1. In the limit where one client is perfectly intelligent of the price ($\epsilon \to 0$), the market maker places a Dirac mass on the order book. In particular, he only trades on profitable sections of the book.

2. In the noise trade limit ($\epsilon \to \infty$), the market maker simply reproduces the client alpha distribution.

3. If the return distribution of the asset is mean-reverting (which is not the case for options with maturities, for example), then so will $\mu_t$ and hence $\gamma_t$ and $c_t$. In the case of a European option, $\gamma''_t$ converges to a Dirac at the payoff at maturity.
Appendix A

Appendix

A.1 A convenient form of the stochastic maximum principle

We present a form of the Pontryagin stochastic maximum principle tailored to the needs of the analysis of section 6.2. The set-up is quite general, following loosely [44]. We assume that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ is a filtered probability space, $(\mathcal{F}_t)_{t \geq 0}$ is generated by a Wiener process $(W_t)_{t \geq 0}$, and we let $A_t(\omega)$ be such that for all $(t, \omega)$, $A_t$ is a Borel convex subset of a Polish topological vector space $E$, adapted to $(\mathcal{F}_t)_{t \geq 0}$. The admissible control processes $(\alpha_t)_{t \geq 0}$ are the progressively measurable processes in $E$ such that $\alpha_t \in A_t$ for all $t \geq 0$. We also assume that the dynamics of the controlled state $X$ are governed by an Ordinary Differential Equation (ODE) with random coefficients valued in $\mathbb{R}^n$:

$$dX_t = f_t(X_t, \alpha_t)dt$$  \hspace{1cm} (A.1)
where the coefficient \( f : (\omega, t, X, a) \to f_t(X, a) \) is Lipschitz in \( X \) uniformly in all the other variables. Assume an estimate of the form

\[
|X_t| \leq |X_0|e^{Ct}
\]  

(A.2)

where \( C \) is a constant. We also assume that \( f \) is continuously differentiable (i.e. \( C^1 \)) in \( X \). The controller’s objective function is given by

\[
J(\alpha) = \mathbb{E} \left[ \int_0^\tau j_t(X_t, \alpha_t)dt + g_\tau(X_\tau) \right]
\]  

(A.3)

where \( \tau \) is a stopping time adapted to \((\mathcal{F}_t)_{t \geq 0}\) and \( j \) and \( g \) are random real-valued concave functions which are \( C^1 \) in \( X \). Furthermore, we assume that \( j \), \( g \) and \( \tau \) are such that \( \mathbb{E} \left[ |g_\tau(X_\tau)| \right] \) and \( \mathbb{E} \left[ \int_0^\tau |j_t(X_t, \alpha_t)|dt \right] \) are finite and uniformly bounded from above. Next, we define the Hamiltonian

\[
\mathcal{H}_t(\omega, X, a, Y) = f_t(\omega, X, a)Y + j_t(\omega, X, a)
\]  

(A.4)

for \( Y \in \mathbb{R}^n \) and for each admissible control \( (\alpha_t)_{t \geq 0} \), the adjoint equation:

\[
-dY_t = \partial_X \mathcal{H}_t(X_t, \alpha_t, Y_t)dt - Z_tdW_t
\]  

(A.5)

with \( Y_\tau = \partial_X g(X_\tau) \) and \( Z_t \) such that \( \mathbb{E} \int_0^\tau |Z_t|^2dt < \infty \). Under these conditions we have the following result.

**Theorem A.1.1** (Pontryagin’s maximum principle). *Let \( \hat{\alpha} \) be an admissible control and \( \hat{X} \) be the associated state variable. Suppose there exists a solution \((Y, Z)\) to the adjoint equation (A.5) such that

\[
\mathbb{E} \left[ \int_0^\tau e^{Ct}|Z_t|^2dt \right] < \infty,
\]  

(A.6)
and the Hamiltonian verifies for every $t \geq 0$

$$H_t(\dot{X}_t, \hat{\alpha}_t, Y_t) = \max_{a \in A_t} H_t(\dot{X}_t, a, Y_t), \quad \text{a.s.} \quad (A.7)$$

and for every $t \geq 0$, almost surely, the function

$$(X, \alpha) \mapsto H_t(\dot{X}_t, \hat{\alpha}_t, Y_t)$$

is concave, then $\hat{\alpha}$ is an optimal control, that is, $J(\alpha) \leq J(\hat{\alpha})$ for all admissible $\alpha = (\alpha_t)_{t \geq 0}$.

**Proof.** If $(X_t)_{t \geq 0}$ is the state associated to another admissible control $(\alpha_t)_{t \geq 0}$, we have the chain of relationships:

$$\begin{align*}
\mathbb{E} \left[ g_\tau(X_\tau) - g_\tau(\dot{X}_\tau) \right] \\
\leq \mathbb{E} \left[ (X_\tau - \dot{X}_\tau) \cdot Y_\tau \right] \\
= -\mathbb{E} \left[ \int_0^\tau (X_t - \dot{X}_t) \cdot \partial_X H_t(\dot{X}_t, \dot{\alpha}_t, Y_t) dt \right] \\
&\quad + \mathbb{E} \left[ \int_0^\tau (f_t(X_t, \alpha_t) - f_t(\dot{X}_t, \dot{\alpha}_t)) \cdot Y_t dt \right] \\
&\leq \mathbb{E} \left[ \int_0^\tau H_t(\dot{X}_t, \dot{\alpha}_t, Y_t) - H_t(X_t, \dot{\alpha}_t, Y_t) dt \right] \\
&\quad + \mathbb{E} \left[ \int_0^\tau (f_t(X_t, \alpha_t) - f_t(\dot{X}_t, \dot{\alpha}_t)) \cdot Y_t dt \right] \\
&\leq \mathbb{E} \left[ \int_0^\tau (H_t(\dot{X}_t, \dot{\alpha}_t, Y_t) - H_t(X_t, \dot{\alpha}_t, Y_t) - H_t(\dot{X}_t, \dot{\alpha}_t, Y_t) + H_t(X_t, \alpha_t, Y_t)) dt \right] \\
&\quad - \mathbb{E} \left[ \int_0^\tau (j_t(X_t, \alpha_t) - j_t(\dot{X}_t, \dot{\alpha}_t)) dt \right] \\
&\leq -\mathbb{E} \left[ \int_0^\tau j_t(X_t, \alpha_t) - j_t(\dot{X}_t, \dot{\alpha}_t) \right]. \quad (A.12)
\end{align*}$$
(A.8) stems from the concavity of \( g \) and terminal condition of \( Y \). (A.9) follows from the dynamics of \( Y \) and the relationship:

\[
\mathbb{E} \left[ \int_0^T \left( X_t - \hat{X}_t \right)^2 Z_t^2 dt \right] \leq 2X_0 \mathbb{E} \left[ \int_0^T e^{Ct} Z_t^2 dt \right] < \infty
\]  

(A.13)

which guarantees that the local martingale part is a martingale. (A.10) holds by the concavity of the Hamiltonian. (A.11) is just the definition of \( \mathcal{H} \). Finally, (A.12) is a consequence of the fact that \( \hat{\alpha}_t \) maximizes \( \mathcal{H}_t(\hat{X}_t, \cdot, Y_t) \).

\[\square\]

### A.2 Derivation of the SPDE (6.43)

We give the main steps of the proof of Proposition 6.3.1. By independence of the common randomness \( W \) and the idiosyncratic randomness \((M^i, \alpha_i^0, \epsilon_i)_{i \geq 1}\), we can freeze the randomness of \( W \) and work after conditioning with respect to \( \mathcal{F}^W \), the \( \sigma \)-algebra generated by \( W \), without affecting the independence properties of the idiosyncratic random variables. By independence of the \( M^i \), the \( \alpha_i^0 \) and \( \epsilon_i \), we have that, conditional on \( \mathcal{F}^W \), the random variables

\[
\alpha_i^t = \alpha_i^0 e^{-\rho t} + \int_0^t e^{-\rho (t-s)} (\nu dW_s + \sigma dM^i_s)
\]

(A.14)

are iid and Gaussian. So if \( f \in C^2 \) is such that \( f \) and its two derivatives have at most polynomial growth, given the assumption (6.44) on the joint distribution \( m \) of all the couples \((\alpha_i^0, \epsilon_i)\), we can apply the law of large numbers and get:

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n \epsilon^i f(\alpha_i^t) = \mathbb{E} \left[ \epsilon^1 f(\alpha_1^t) \right] | \mathcal{F}^W
\]

\[
= \int_{\mathbb{R}^3} \epsilon f \left( \alpha_0 e^{-\rho t} + \int_0^t e^{-\rho (t-s)} \nu dW_s + \sigma x \sqrt{\int_0^t e^{-2\rho (t-s)} ds} \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} m(d\alpha_0, d\epsilon) dx
\]
The other results can be derived directly from this explicit representation.
Bibliography


