Department of Political Economy

How to Measure Market Liquidity Risk in Financial Institutions?

Joana Maria Cabrita Lopes

A Dissertation presented in partial fulfillment of the Requirements for the Degree of
Master

Supervisor:
Dr. João Pedro Pereira, Assistant Professor,
ISCTE Business School – Lisbon University Institute

Co-supervisor:
Dr. Umut Çetin, Lecturer in Statistics,
Department of Statistics – London School of Economics and Political Science

September, 2010

"*Liquidity always comes first; without it, a bank doesn't open its doors; with it, a bank may have time to solve its basic problems.*"

Chief Financial Officer at Citigroup

Summary

We apply numerical stochastic dynamic programming to derive trading strategies that minimize the mean and variance of the costs of executing a large block of a security over a fixed exogenously defined time period. Financial markets are considered to be liquid if a large quantity can be traded quickly and with minimal price impact. Although, the trading costs associated with trading such large quantity of a single asset – often called execution or transaction costs – can be substantial significant that directly influence the return of the investment. To minimize the price impact, an investor would choose to split his order into many small pieces. However the time taken to transact introduces a risk component in execution costs that arise from unfavourable price movements during the execution of an order. The longer the trade duration, the higher the uncertainty of the realized prices. In this setting, the decision can be viewed as a risk/reward trade-off faced by the investor who not only cares about the expected value but also about the variance (or volatility) of his execution costs. Risk aversion in this context means that an investor is willing to trade lower risk for higher price impact costs. A numerical solution for minimizing a combination of the expected transaction costs and volatility (or price) risk is derived. The parameters of the price impact model are estimated based on real world stock data.

Keywords: market liquidity risk; transaction costs; optimal trading strategies; stochastic dynamic programming.

JEL classification: G12 and G32

Resumo

A globalização e o desenvolvimento dos mercados financeiros nos últimos anos implicaram uma crescente dependência das instituições financeiras do financiamento nos mercados internacionais, com a utilização de instrumentos financeiros cada vez mais complexos, criando assim novos desafios na gestão do risco de liquidez. Este desenvolvimento dos mercados e a recente crise de 2008 realçaram a importância vital de existir um adequado sistema de mensuração do risco de liquidez para uma melhor eficácia no funcionamento do sector bancário.

No período que antecedeu a crise do *subprime*, os mercados estavam confiantes e o financiamento estava facilmente acessível e a baixo custo. A alteração das condições de mercado ilustraram quão rapidamente a liquidez se pode evaporar e repercutir-se durante um longo período.

O caso LTCM (*Long Term Capital Management*) tem um especial interesse para a gestão do risco de liquidez uma vez que as posições detidas pelo fundo revelaram ser demasiado elevadas para serem liquidadas sem induzir grandes movimentos nos preços de mercado devido à insuficiente liquidez do mesmo. A escassez repentina de liquidez nos mercados é um sintoma observado na maioria das crises financeiras. Assim, a identificação, quantificação, monitorização e controlo do risco de liquidez assumem um papel de destaque quer para as instituições financeiras quer para os reguladores.

Os modelos usados na quantificação do risco de mercado, tipicamente *Value at Risk models*, geralmente não consideram se o preço de mercado de um determinado título (ou carteira de títulos) pode ou não ser realizado em caso de liquidação, ou seja não consideram o risco de liquidez de mercado. Esta situação pode conduzir a uma sub-estimação do risco total e consequentemente a uma errada alocação de capital.

Neste sentido, esta dissertação pretende responder à questão: Como medir o risco de liquidez (de mercado) nas instituições financeiras?

Podemos distinguir dois principais tipos de risco de liquidez: o risco de liquidez de financiamento e o risco de liquidez de mercado, existindo uma forte relação entre ambos. O risco de liquidez de financiamento é o risco de um banco não poder honrar os

seus compromissos financeiros nas datas devidas sem incorrer em perdas significativas. A consequente necessidade de financiamento pode requerer a venda de activos podendo afectar a liquidez de mercado. O risco de liquidez de mercado é o risco de uma posição não poder ser facilmente liquidada (e num curto espaço de tempo) sem influenciar substancialmente o preço de mercado.

Apesar da ligação entre os dois tipos de risco, estes são objecto de estudo de áreas distintas da economia e finanças. O primeiro é estudado no âmbito da Gestão de Activos e Passivos (ALM) e o segundo, o risco de liquidez de mercado, é um tópico da micro-estrutura dos mercados e das estratégias óptimas de negociação.

As estratégias óptimas de negociação dizem respeito à gestão e mensuração dos custos associados à transacção de títulos e à definição de estratégias que minimizam esses custos. Assim, medir o risco de liquidez de mercado implica medir os custos de negociação, que, embora incertos, dependem do impacto no preço o qual é influenciado pelo volume transaccionado. O risco de liquidez de mercado, e consequentemente as estratégias óptimas de negociação, serão o tema central desta dissertação.

Um problema típico enfrentado pelas instituições financeiras (e pelos grandes investidores, e.g., os investidores institucionais) é a liquidação (ou aquisição) de grandes posições num determinado activo, tal como um grande volume de acções. Considera-se que os mercados financeiros são líquidos quando uma grande quantidade pode ser transaccionada rapidamente e com um impacto mínimo no preço. No entanto, a execução imediata frequentemente não é possível ou apenas é a um custo demasiado elevado devido à reduzida liquidez do mercado.

O impacto no preço e os custos de execução (também denominados de custos de transacção ou de negociação) podem ser significativamente reduzidos dividindo a ordem (de venda ou de compra) em ordens mais pequenas repartidas por um determinado horizonte temporal. Assim, uma questão pertinente é: como definir estratégias óptimas de negociação de modo a que os custos esperados de execução sejam minimizados? Problemas deste tipo têm sido objecto de estudo de vários autores, entre os quais se destacam Bertsimas and Lo (1998).

Contudo, o tempo total necessário para executar uma grande quantidade introduz uma componente de risco nos custos de execução que resulta dos movimentos não favoráveis no preço que podem ocorrer durante o período de execução. Quanto maior o tempo de execução maior será a incerteza dos preços realizados. Os investidores avessos ao risco negociarão mais rápido, incorrendo em custos de transacção mais elevados mas com menor risco. Neste sentido, a decisão pode ser encarada como um custo/beneficio do investidor que tem em conta não somente o valor esperado dos custos mas também a variância dos mesmos.

Assim, considerando apenas o custo esperado de execução como 'função objectivo' deixa de parte uma importante componente da liquidez que é o risco de volatilidade que está associado ao prolongar (a venda ou compra) de uma transacção. Por este motivo, Almgren and Chriss (2000) sugeriram substituir a minimização dos custos esperados pela minimização do valor esperado e da variância dos custos resolvendo o respectivo problema de optimização na classe das estratégias determinísticas (ou estáticas).

No entanto, o simples acto de negociar afecta não só os preços actuais mas também a dinâmica de preços, que por sua vez, afecta os custos de negociação futuros. Assim, medir os custos de transacção é um problema fundamentalmente dinâmico e não estático.

Por consequência, estudou-se o modelo dinâmico de Bertsimas and Lo alterando a 'função objectivo' de modo a incorporar o risco de volatilidade (ou risco de preço). Em vez de se minimizar apenas os custos esperados de executar um grande volume de acções durante um período finito de tempo (exogenamente definido) derivou-se uma estratégia óptima de negociação que minimiza uma combinação entre os custos de transacção e o risco. Este problema de optimização pode ser resolvido recorrendo à programação dinâmica estocástica e resolvido numericamente à luz da equação de Bellman (1957). Sendo um problema recursivo o algoritmo utilizado foi o algoritmo de indução inversa, ou seja, indução do futuro para o presente (*backward induction*), uma vez que no último período o número de acções a negociar é conhecido (são as que restam).

No modelo de Bertsimas and Lo o preço de execução é composto por duas componentes, uma componente sem impacto no preço, que resulta da evolução normal do preço na ausência de impacto (pode ser medida pelo ponto médio entre o preço de compra e venda), e uma componente denominada 'impacto no preço' que é uma função linear do volume negociado e das condições de mercado (e informação disponível). Os parâmetros do modelo foram estimados com base em dados históricos de bolsa. Na estimação dos parâmetros da componente ‘impacto no preço’ utilizou-se uma regressão linear.

Com base no algoritmo de optimização, desenvolvido em linguagem MATLAB, fez-se uma análise comparativa entre a estratégia óptima de negociação que considera a componente da volatilidade (ou risco) e a que não considera, para diversos valores dos parâmetros. Com base nos resultados as principais conclusões foram as seguintes:

- O risco (caracterizado por uma função objectivo quadrática) é uma componente importante dos custos de transacção que não deve ser ignorada;
- Os custos de execução aumentam com a quantidade transaccionada, ou seja, quanto maior for o volume transaccionado maior será o impacto no preço e consequentemente maiores serão os custos de transacção;
- Quando o peso da informação disponível aumenta os custos de transacção diminuem, uma vez que o acesso à informação e às condições de mercado implicam um conhecimento da tendência dos preços de mercado podendo o investidor tirar partido dessa informação;
- Existem evidências que levam a concluir que o aumento da volatilidade do título negociado aumenta os custos de transacção;
- Quanto maior o tempo total da transacção menor serão os custos de execução, dado que a quantidade transaccionada vai diminuindo;
- Os investidores mais avessos ao risco assumem maiores custos de transacção de modo a reduzirem a sua exposição ao risco.

Acknowledgements

Firstly, I would like to thank Professor João Pedro Pereira for his availability, awareness and background, and for his valuable comments.

As well, I would like to express my gratitude to Professor Umut Çetin for his thoughts and for giving me the opportunity to live an unforgettable experience at the London School of Economics and Political Science.

To Pedro for always believing in me and for his encouraging words in the troubling moments.

And last but not least to my sister Margarida and to my parents Esilda and José for their unconditional support.

Contents

List of Tables

List of Figures

List of Abbreviations

ALM	Asset and Liability Management
BLUE	Best Linear Unbiased Estimator
DP	Dynamic Programming
DW	Durbin Watson (Test)
EC	Execution Costs
EWMA	Exponentially Weighted Moving Average
HJB	Hamilton-Jacobi-Bellman (Equation)
IID	Independently and Identically Distributed
LPT	Linear-Percentage Temporary
LTCM	Long Term Capital Management
LTL	Lithuanian Litas
LVaR	Liquidity-adjusted Value at Risk
MTM	Mark-To-Market
OLS	Ordinary Least Squares
PDF	Probability Density Function
USD	United States Dollar
VaR	Value at Risk
WN	White Noise

Chapter 1

Introduction

Financial market developments in the last years, such as the increasing reliance of large institutions on market funding, the increasing use of complex financial instruments, and the globalisation of financial markets, have created significant new challenges in liquidity risk management[1]. These market developments, and the 2007-2008 market turmoil, highlight the vital importance for the soundness of the banking sector to have adequate liquidity risk measurement systems for both normal and stressed times, and to maintain adequate liquidity buffers.

Prior to the turmoil, markets were confidant and funding was readily available at low cost. The change in market conditions illustrated how quickly liquidity[2] can evaporate and that illiquidity can last for an extensive period of time.

The case of Long Term Capital Management (LTCM) is also of special interest for liquidity risk management because positions held by the fund were too large to be liquidated without inducing major price movements due to insufficient market liquidity. Commonly, the sudden dry up of market liquidity is a symptom observed in most of financial market crises.

The Value at Risk (VaR) models, often used for the estimation of market risk, generally do not consider whether the market price of a security can actually be realized in case of a liquidation. This may lead to an underestimation of the total risk, and hence to a misleading capital allocation.

[1] The fundamental role of banks in transforming short-term deposits into long-term loans makes banks inherently vulnerable to liquidity risk. Liquidity risk management is the constant process of balancing the cash inflows and outflows from on- and off-balance sheet items, along with structural and strategic planning, to ensure both that adequate sources of funding are available, and that those sources are used properly. Liquidity risk management also requires robust internal governance, adequate tools to identify, measure, monitor, and control liquidity risk, including stress tests and contingency funding plans.

[2] Liquidity, in the broadest sense of the term, is the capacity to obtain funding when it is needed.

Therefore, this dissertation aims to answer the question: How to measure (market) liquidity risk in financial institutions?

One can distinguish two main types of liquidity risk, funding liquidity risk and market liquidity risk[3]. Funding liquidity risk is the risk that a bank will not be able to honour its financial commitments when they are due without incurring substantial loss. Market liquidity risk is the risk that a position cannot easily be liquidated without significantly influencing the market price.

A typical problem faced by financial institutions is the liquidation (or acquisition) of a large asset position, such as a large block of shares. An immediate execution is often not possible or only at a very high cost due to a scarce liquidity of the market.

The overall price impact and the execution costs (also called 'transaction costs') can be significantly reduced by splitting the order into a sequence of smaller orders that are spread over a certain time horizon. Hence, one pertinent issue is to find optimal trading strategies such that the expected execution costs are minimized. Problems of this type were analyzed by many authors, namely Bertsimas and Lo (1998).

Nevertheless, the time taken to transact introduces a risk component in execution costs. Risk averse agents will trade more rapidly thus incurring higher transaction costs but lower risk. Consequently, taking the expected execution costs as a target function misses an important component of liquidity, the volatility risk that is associated with delaying an order. For that reason, Almgren and Chriss (2000) suggested replacing the minimization of expected costs by a mean-variance optimization of costs solving the correspondent optimization problem in the class of deterministic (or static) strategies.

Although measure executions costs is a dynamic (stochastic) problem, not static, since trading transactions affects both current and future prices.

We therefore propose to study the 'linear-percentage temporary' dynamic model of Bertsimas and Lo by changing the objective function in order to incorporate the volatility risk component (or price risk). Instead of minimizing merely the expected transaction costs of trading a large block of equities over a (fixed) finite horizon we derive a dynamic optimal trading strategy that minimizes a combination of trading costs

[3] Furthermore, in literature, we can find another liquidity risk types, such as: call liquidity risk, term liquidity risk and contingent liquidity risk. For more detail please see Appendix A.

and volatility (or price) risk. It can be seen as a typical problem of stochastic dynamic programming and can be solved numerically based on the 'Bellman's Equation'.

This dissertation is organized as follows: Chapter 2 introduces some concepts related with market liquidity risk and his components; Chapter 3 discusses optimal trading strategies. In Chapter 4 we review dynamic programming theory. Chapter 5 explicitly examines the linear percentage temporary model and Chapter 6 presents and discusses numerical results based on real data. Last chapter concludes.

Chapter 2

Market Liquidity

Asset returns are usually calculated using mid market or closing prices. Typical measures of market risk are based on these returns. Implicitly, it is assumed that these are the prices that can be achieved in case of liquidation. This is the point, where (market) liquidity risk comes into play. Market liquidity risk deals with the risk of losses arising from the deviation of the realised price in a buying/selling process as compared to the market price prevailing prior to the transaction. This loss is denoted as the transaction cost, representing an additional charge when buying, and taking the form of a price discount when selling an asset. The relevant market price is the mid-price between bid and ask as the best available estimate of the fair value of the security. By taking the mid-price, trading cost is split equally between buy and sell transactions.

2.1 Market Liquidity and Funding Liquidity: Definition and Interactions

Funding (or cash flow) liquidity risk is the risk that a bank will not be able to honour its financial commitments (both expected and unexpected current and future cash flows) when they are due without incurring substantial loss. The consequent need for cash may require selling assets. Market (or asset) liquidity risk is the risk that a position cannot easily be liquidated at short notice without significantly influencing the market price, because of inadequate market depth or market disruption, and hence the liquidation value of the position will differ significantly from their current mark-to-market (MTM) value[4]. Thus liquidity risk can arise from both, the assets and liabilities of a financial institution.

[4] Mark-to-market value refers to accounting for the value of an asset or liability based on the current market price of the asset or liability.

The increasing market-based funding of banks originate a correlation between funding liquidity risk and market liquidity risk since market illiquidity can difficult a bank to raise funds by selling assets and thus increase the need for funding liquidity. The resulting changes in demand for funds can, afterwards, affect market liquidity.

Despite the relation between market and funding liquidity, usually they are treated in two distinct branches of economics and finance.

Funding liquidity risk results from size and maturity mismatches of assets and liabilities which is subject of Asset and Liability Management (ALM), while market liquidity risk is a topic of market microstructure theory and optimal trading strategies. Therefore, concepts for measuring and managing the two types of liquidity differ substantially from each other.

Assessing funding liquidity risk implies checking the asset and liability structure of a financial institution and the potential demands on cash and other sources of liquidity.
Asset and Liability Management unit is in charge of managing the differences, at all future dates, between assets and liabilities of the banking portfolio. Controlling (funding) liquidity risk implies controlling over time the cash flows, avoiding unexpected market funding and maintaining a ‘cushion’ of liquid (short-term) assets[5], so that selling them provides liquidity without incurring in losses. Liquidity risk exists when there are deficits of those funds. When there are excess of funds the result is interest rate risk, the risk of not knowing in advance the rate of lending or investing these funds.

Market microstructure theory studies the role of trading mechanisms on the price setting process. This area of literature examines the ways in which the market structure and trading mechanisms, i.e., the evolution process of a market, affects transactions costs, prices, volume and trading behaviour. Optimal trading strategies concerns with the measurement and management of trading costs and the definition of strategies that minimize those costs. Therefore, measuring market liquidity risk implies measure

[5] Liquid assets are usually defined as assets that can be quickly and easily converted into cash in the market at a reasonable cost.

transaction costs which are uncertain and a function of the price impact of trades and the size of the positions. Market liquidity risk (and optimal trading strategies) constitutes the main focus of this dissertation and will be treated in depth in subsequent sections.

2.2 Trading Costs Components

Trading costs, also called execution or transaction costs, have several components: specific costs such as commissions and bid-ask spreads, and costs that are harder to quantify, such as the opportunity cost of waiting and the price impact from trading. Opportunity costs arise because market prices are moving constantly and can change favourably or unfavourably, generating unexpected profits or lost opportunities while a trader hesitates.

Glantz and Kissell (2003) identify nine components of trading costs: broker commissions, exchange fees, taxes, bid-ask spreads, investment delay, price appreciation, price impact, timing risk and opportunity cost.

2.3 Market Liquidity Measures

Liquidity risk is one of the factors, typically ignored in Value at Risk estimates, which is a widely used measure of market risk. It is assumed that any portfolio position is liquidated in a single block and the mark-to-market value is always fully realised. Since the postulate of infinitely elastic markets contradicts the premise of prudence in risk management, some ad-hoc adjustments to VaR have been proposed. Time horizon and volatility are the two parameters through which the VaR number can be changed in order to account for an increase in liquidity risk. Often, the time adjustment is implemented for all portfolio positions as a ten day holding period which is required by regulatory authorities. Increasing the time horizon over which VaR is calculated (frequently through the “square root of time” rule) to account for the time taken to liquidate a large position, from an economic perspective, this procedure has severe limitations, as capital would be tied up inefficiently if positions can be sold quicker than the assumed minimum holding period. In addition, the “square root of time” rule, assumes that no autocorrelation exists between returns from one measurement period to another. Another possibility to take liquidity risk into account is to multiply the VaR measure by some conservative factor. Nevertheless, all above mentioned approaches leave the question of how much and by what criteria the VaR measure should be altered or not. As so, a liquidity risk measure should be based on quantitative, measurable criteria.

2.3.1 Bid-Ask Spreads

One of the components of transaction costs and most popular measure of liquidity is the bid-ask spread. Lower transaction costs and hence a narrower spread reflects a better liquidity in the market. The absolute bid-ask spread is simply defined as the difference between the (best) ask price (lowest price for which a seller is willing to sell a security) and the (best) bid price (highest price that a buyer is willing to pay for a security). In relative terms can be calculated by relating the absolute spread to the mid price of the security:

$$S_t^{Relative} = \frac{P_t^{ask} - P_t^{bid}}{P_t^{mid}} \qquad (1)$$

where P_t^{bid} and P_t^{ask} are the bid and ask prices of a security in time t and the mid price of a security is defined as:

$$P_t^{mid} = \frac{P_t^{ask} + P_t^{bid}}{2} \qquad (2)$$

Relative spreads enhance comparability between the spread sizes of different securities.

Bangia et al. (1999) further split liquidity risk into an exogenous and an endogenous component. Exogenous illiquidity is determined by factors beyond the individual trader's control and is equally relevant for all market participants. It describes the part of liquidity cost which is not affected by the size of the position in the market, is the result of market characteristics. Observable variables, such as depth and bid-ask spread, make it measurable. In contrast, endogenous liquidity risk is specific to the position in the market and varies across market participants. It is mainly driven by the size of a position held: the larger the size of the position, the greater is the exposure to endogenous liquidity risk. Nevertheless, it can be influenced by the trader's own actions by applying appropriate trading schedules like splitting a large order into smaller pieces.

Bangia et al. (1999) only include exogenous liquidity risk in their liquidity adjusted VaR measure considering the uncertainty in the spread. They characterize the distribution of the relative spread by its mean ($\bar{S}$) and standard deviation or volatility ($\tilde{\sigma}$). They adjust the VaR measure considering the worst increase in the spread at some confidence level, known as Liquidity-adjusted Value at Risk (LVaR) which combines market and liquidity risk:

$$LVaR_{(1-\alpha)\%} = VaR_{(1-\alpha)\%} + \frac{1}{2}[P_t^{mid}(\bar{S} + a\tilde{\sigma})] \qquad (3)$$

where a is the scaling factor such that (1-α)% percent of liquidation cost is covered.
This assumes that the worst market loss and increase in spread will occur at the same time. In general is true, we observe a correlation between returns and spreads.

Although this approach has the merit of considering some transaction costs, it only looks at the bid-ask spread component of this costs, which may be enough for a small position of a stock but is not when liquidation can affect market prices. The price impact factor should be taken into account.

2.3.2 Price Impact

The liquidation price is not only a function of time but also of position size. For positions up to a specific size, usually the current market depth, the transaction can be accomplished at the bid or ask price depending on the direction of the trade. For quantities exceeding market depth, an additional price discount, commonly called price (or market) impact, has to be accounted for. Price impact is the typically unfavourable effect on prices that the process of trading creates: a security's seller will, by the very act of selling, push down the security's price, and, in the opposite way, a security's buyer will, by the very act of purchasing, push up the security's price.
Moreover, the sell/buy price is a decreasing/increasing function of quantity and the larger the order, the more heavily the trade affects the price. However, the functional form that this price-volume relationship takes is not clear. For simplicity, it is often assumed to be linear, whereas there is empirical evidence of concave in case of a buy order as well as convex in case of a sell order. This price-size function is illustrated in

Figure 1, which shows the relationship between the liquidation price and the total position size held.

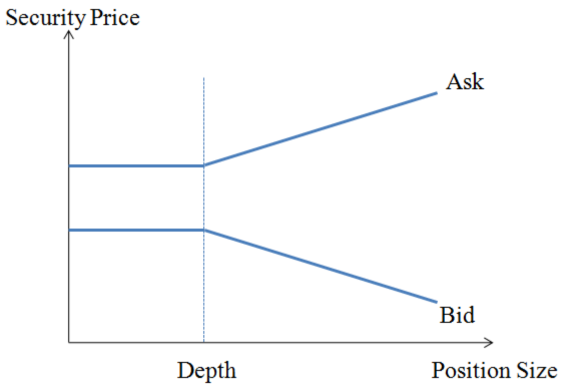

Figure 1 – Effect of position size on security price

Price impact is an important parameter for the estimation of optimal liquidation strategies, as it can be reduced by slicing the order into several trades. The premium paid for a buy order and the concession made for a sell order can be interpreted as incentives for other market participants to provide additional liquidity. Thus, a large order can be completed within reasonable time.

Bertsimas and Lo (1998) proposes a model which, given a price-impact function, furnishes the optimal sequence of trades that minimizes the expected transaction costs. Nevertheless, their approach ignores the volatility of costs for different trading strategies, i.e., a penalty for the uncertainty of cost (or revenue).

2.3.3 Expected Transaction Costs and Volatility

The drawback of liquidating a large block of a security more slowly, however, is that the portfolio remains exposed to price risk over a longer period. The immediate sale yields to a high cost but minimum risk. Under the uniform sale, the position is sold off in equal size blocks, leading to low cost but higher volatility. Execution strategies do not need to be limited to those two extreme cases – immediate or uniform liquidation. More generally, we can select a strategy that leads to an optimal trade-off between execution costs and price risk.

Almgren and Chriss (2000) examined market liquidity risk by minimising a combination of volatility risk and expected transaction costs arising from (permanent and temporary[6]) price impact. The rationale is the following: a trading schedule is only optimal if it involves least cost for a certain level of risk and also least risk for a certain level of cost. Such strategies can be determined by minimising the mean-variance objective function for various levels of risk aversion.

They considered the risk-reward trade-off both from the point of view of classic mean-variance optimisation and of VaR. Their analysis led to general insights into optimal trading strategies, and to several applications including a definition of liquidity-adjusted VaR.

2.4 Market Liquidity Features

Market liquidity is usually determined by four key factors:

The *tightness* of the market, which is measured using the bid-ask spread (as seen before, the difference between buy and sell prices), determines the cost of unwinding a position at short notice. The smaller the difference between the ask and bid prices, the better the liquidity in the market.

The *depth* of the market assesses which transaction volume can be realised immediately without affecting prices. Small amounts should be able to be traded without impact on prices, for large amounts, a premium for buy orders and a discount for sell orders have to be accepted.

Resiliency describes the speed at which market prices return to equilibrium after a major transaction.

Immediacy is defined as the time between the start of a market transaction and its final completion. It denotes the speed with which a position is liquidated.

If demand meets supply, even for relatively large trading volumes, and if price impact is minimal, transaction costs are low and the market is considered liquid. A market is

[6] The difference between temporary and permanent price impact will be described afterwards in Chapter 3.

perfectly liquid if any volume can be traded at any time at no cost. Since there are no infinitely deep markets, an investor who wishes to trade a large block immediately, as we mentioned before, needs to pay a premium for a buy order or accept a discount for a sell order. These factors increase with the size of the transaction. By splitting the block into smaller orders the investor should be able to reduce transaction costs. The splitting strategies which minimize the transaction costs are known as "optimal liquidation strategies" and would be explained in the subsequently chapter. Another way to minimize transaction costs is to limit the exposure of the security in order to avoid a large price impact in case of forced liquidation.

Chapter 3

Optimal Liquidation Strategies

An optimal trading strategy describes how a large order should be sliced into pieces over a period of time. If a large position is held, the financial institution most likely won't liquidate it all at once, but will rather split it up into several orders. This strategy reduces expected transaction costs, implying that liquidation costs and risks will depend on the strategy and the time horizon chosen. Trying to find an optimal trading strategy, not only price impact, but also volatility risk has to be taken into account. During the liquidation period, remaining parts of the holdings are exposed to price risk. An optimal liquidation strategy is the result of a balance between a reduction in potential price impact and an additional exposure to volatility (or price) risk.

The priority objective for the design of such a strategy is to preserve asset value, that is, to minimise the cost component in the presence of risk. In its basic form, the optimization condition only involves price impact as the cost component and price volatility as the only risk factor. A strategy is evaluated according to its expected cost-risk profile: an aggressive strategy, characterised by large initial trades, leads to high price impact cost, but reduces price risk by quickly selling off or buying the remaining shares. Trading more passively, namely shifting parts of the trades to later periods would cause less price impact at the expense of an increase in risk. As price impact costs are a decreasing function of time and the risk of a strategy is increasing over time, it can, as a general rule, be concluded that the change of one term affects the other adversely.

3.1 Temporary and Permanent Price Impact

Some authors distinguish two kinds of price impact: temporary and permanent. Temporary impact refers to temporary imbalances between supply and demand caused by our trading leading to temporary price movements away from equilibrium. Permanent impact means changes in the "equilibrium" price due to our trading, which remain at least for the life of our liquidation. Figure 2 represents the price movement during the trading period.

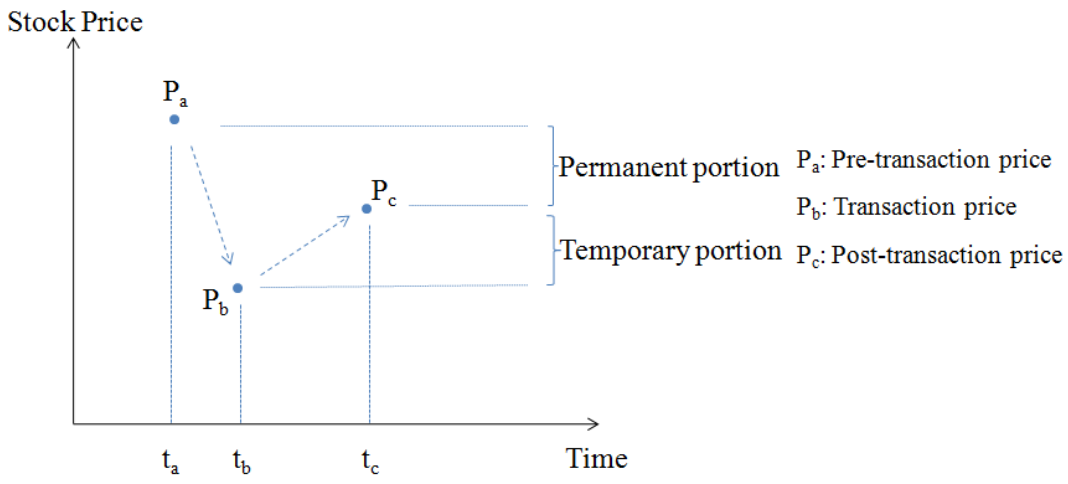

Figure 2 – Permanent and temporary price impact from sale

3.1.1 Temporary Price Impact

Temporary price impact arises due to a short-term demand and supply imbalance caused by one's own order. When a trader posts a sell order exceeding market depth, he has to accept lower prices in order to complete the trade. If the motives of the trader were purely liquidity-based and other market participants were aware of this motivation, the price would fully recover shortly after the trade hits the market. One of the key questions is how long it takes for the price to return to its pre-trade level. The most common assumption is that temporary price impact will have dissipated completely until the next trade. This implies that in such a model trading intervals cannot be chosen arbitrarily small, as prices need time to return to equilibrium.

The larger the order one wishes to trade in a period, the higher will be the premium the market requires for a buy order (or the concession for a sell). Therefore, temporary price impact in its simplest form is chosen to be an increasing linear function of trading size in a particular period.

Almgren and Chriss (2000), however, state that the true price impact function is probably nonlinear. Hence, in the later papers Almgren (2003) and Almgren (2005) extend the price impact model to accommodate nonlinear functions for temporary price impact.

3.1.2 Permanent Price Impact

Market participants, observing a large trade in the market, might infer that the buyer/seller is in possession of some private information. As a consequence, they adjust their beliefs about future prices. This way, a permanent change in the equilibrium price is induced by one's own trading. Prices don't return to their original trajectory after a trade, but rather follow a new one that better reflects the true market value. Therefore in contrast to temporary price impact, which dissipates quickly, price dynamics have to be adjusted for the permanent impact factor. This permanent impact factor is usually assumed to be linear in total transaction size. Contrary to temporary price impact, permanent impact may not be influenced through the speed of trading because it depends on total transaction size and not on the size of the trade in a particular period.

3.2 Static and Dynamic Strategies

Broadly speaking we can distinguish two types of trading strategies: dynamic and static. Static strategies are determined in advance of trading, that is the rule for determining each piece size of the sliced trading depends only on information available at the starting time t=1. Dynamic strategies, conversely, depend on all information up to and including time T-1.

3.3 Extreme Strategies

In trading a highly illiquid, volatile security, there are two extreme strategies, as mentioned before: trade the whole position immediately at a known, but high cost, or trade in equal sized packets over a fixed time at relatively lower cost. The latter strategy corresponds to trading at a constant rate and has lower expected costs but this strategy has the disadvantage of greater uncertainty in final cost (the variance may be large if the period is too long). How to evaluate this uncertainty is subjective but also a function of the trader's tolerance for risk. What we know is that for a given level of uncertainty, cost can be minimized.

The strategy at the other end of the scale, where the whole position is sold off in the first period, minimises variance (has the smallest possible variance). Here, the price impact is the highest, compared to all other optimal strategies. This means, a risk-averse trader always favours to liquidate large parts of the total transaction early. The different strategies are compared in Figure 3, we can observe that order sizes are declining over the liquidation horizon:

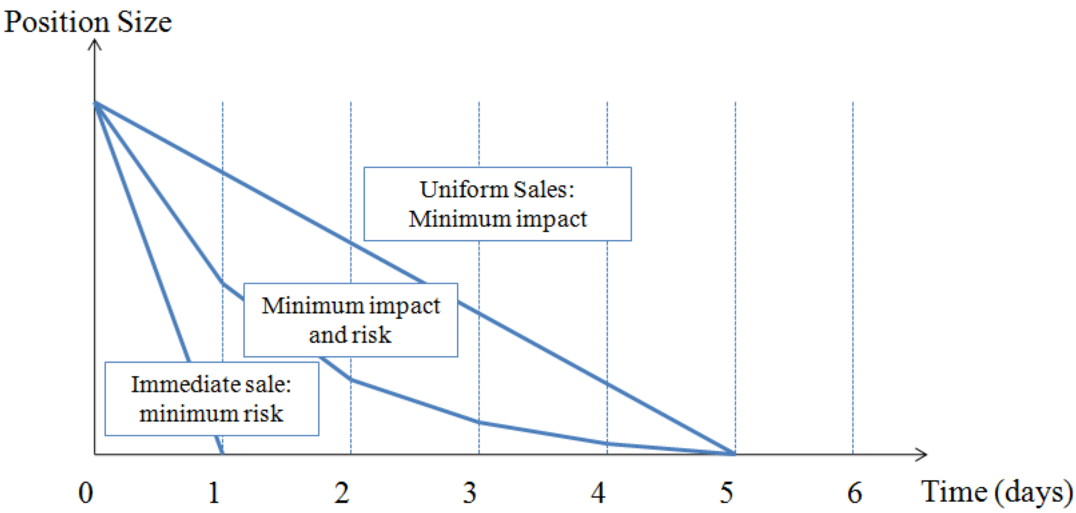

Figure 3 – Different execution strategies

The purpose of this thesis is to show how to compute optimal strategies that lie between these two extremes.

3.4 Implementation Shortfall

Suppose that the initial security price is S_1 and the total trade size N, so that the initial market value of our position is $N \times S_1$. The security's price evolves according to two exogenous factors: volatility and drift, and one endogenous factor: price impact. Volatility and drift are assumed to be the result of market forces that occur randomly and independently of our trading.

The total cost of trading is the difference between the initial book value $N \times S_1$ and the realized book value $\sum_{t=1}^{T} n_t S_t$, where $\sum_{t=1}^{T} n_t = N$. This is the standard ex-post measure of transaction costs that will be used in this dissertation to evaluate transaction costs, and is essentially what Perold (1988) calls 'implementation shortfall'.

3.5 Application of Optimal Trading Strategies in Liquidity Risk Management

Trading cost at some future point in time is uncertain, because market conditions change over time. The difference between the realised liquidation price and the market price is determined by the supply and demand curve. Especially, in case of financial market crises, it is important to bear this uncertainty in mind. Under such circumstances, liquidity often dries up quickly and positions can only be unwound by taking much larger losses than usual. In order to prepare for such scenarios, stress tests, making worst case assumptions and considering the potential rise in bid-ask spreads, should be conducted.

Additionally, optimal trading strategies could be adapted to be used in stress-testing. They can be either useful for simulating a funding or market liquidity crisis. An advantageous feature is the explicit inclusion of the time component, so that questions essential for funding liquidity management might be addressed. For instance, the time needed to raise a certain amount of cash through the selling of securities could be assessed employing the optimal liquidation theory.

Insights gained from optimal liquidation strategies could likewise be applied for stress testing in market and liquidity risk management[7] by adapting relevant parameters to crisis situations.

The most important variable is probably price impact, which is likely to increase sharply when the market crashes. However, other essential factors should not be neglected: for instance, resiliency is certainly affected, so that prices don't revert quickly to pre-trade levels in case of short liquidity in the market. Therefore, the assumption that temporary price impact dissipates quickly doesn't hold anymore. During normal market times, it could be argued that the time between trading should be lengthened in order to profit better from price reversions. However, the high price volatility in crisis situation doesn't allow for long waiting times.

[7] Liquidity management is the continuous process of raising new funds, in case of a deficit, or investing excess resources when there are excesses of funds.

In order to reduce the risk of large losses due to price impact, firms enforce position limits to traders to limit the exposure to a single instrument.

Note that some portfolios of financial institutions are subject to reserves, which are pricing changes in the valuation away from fair value to account for such effects as illiquidity and model risk. This reserve is deducted from the fair value of positions to account for the time and costs required to close out the position depending on the liquidity of the market. In such cases, there is no need to take liquidity risk into account because it is already reflected into the valuation of positions.

Chapter 4

Dynamic Programming

The optimization of problems over time arises in many settings of the real world, ranging from the control of landing aircraft to managing entire economies. These problems involve making optimal decisions, then observing information, after which we make more decisions, and then observe more information, and so on. Known as dynamic optimization problems or, also called optimal control or sequential decision problems. Dynamic Programming (DP) is a recursive method for solving sequential decision problems.

Dynamic Programming is one of the most fundamental building blocks of modern macroeconomics. It gives us the tools and techniques to analyse (usually numerically but often analytically) a whole class of models in which the problems faced by economic agents have a recursive nature.

The term dynamic programming was first introduced by Richard Bellman, who today is considered as the inventor of this method, because he was the first to recognize the common structure underlying most sequential decision problems.

Dynamic Programming can be a useful algorithmic for deterministic problems, but it is often an essential tool for stochastic problems, which involves uncertainty. It can be applied in both discrete and continuous time settings. In any DP problem, there are two key variables: a state variable and a control variable (or decision variable). The optimal decision is a function dependent on the state variable and time. If the last time period is finite we have a finite horizon problem, if time goes up to infinite we call an infinite horizon problem.

Time	Discrete
	Continuous
Horizon	Finite
	Infinite
Randomness	Deterministic
	Stochastic
State space/Variables	Discrete
	Continuous

Table 1 – Characteristics of a Dynamic Programming

The key idea behind Dynamic Programming is the Principle of Optimality formulated by Richard Bellman (1957):

"*An optimal policy has the property that, whatever the initial state and initial decision (i.e., control) are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.*"

This means that if the part of a control variable from time 0 to T is an optimal program as evaluated at time 0, then at any later time t the same path from t to T must be an optimal program, "in its own right", as evaluated at t.

The solution of a dynamic programming problem traditionally involves backward iteration on the *Bellman's Equation*. Typically, the *Bellman's Equation* cannot be characterized in closed form and therefore has to be approximated. Different approaches have been advanced for *Bellman's Equation* approximation, such as discretization of the state space or projection methods. The state space discretization approach is subject to the 'curse of dimensionality'[8] and, thus, highly inefficient for problems with multiple state variables.

4.1 Value Function

All dynamic programs can be written in terms of a recursion that relates the value of being in a particular state at one point in time to the value of the states that we are

[8] *Curse of dimensionality* is the problem caused by the exponential increase in volume associated with adding extra dimensions to a (mathematical) space. The term was conceived by Richard Bellman.

carried into at the next point in time, called value function. For stochastic problems, this equation can be written as[9]:

$$V_t(x_t) = \max_{c_t}\{u(c_t, x_t) + E_t[V_{t+1}(x_{t+1})]\} \qquad (4)$$

where x_{t+1} is the state we transition to if we are currently in state x_t and take action c_t (control variable). This equation is also known as stochastic *Bellman's Equation*, or the stochastic 'Hamilton-Jacobi-Bellman Equation' (HJB Equation) for continuous problems. Some textbooks (in control theory) refer to them as the functional equation of dynamic programming.

$E_t(y)$ denotes the mathematical expectation of a random variable y, given information known at t. At time t, x_t is assumed to be known, but $x_{t+j}, j \geq 1$is unknown at t.

Note that the maximization (or minimization) over the whole path $c_{1:T}$ has reduced to a sequence of maximizations over c_t. This simplification is due to the Markovian nature of the problem: the future depends on the past, and vice versa, only through the present.

4.2 Numerical Dynamic Programming

The need for a numerical solution arises from the fact that generally dynamic programming problems do not possess tractable closed form solutions. Hence, techniques must be used to approximate the solutions of these problems.

For each problem, specification of the state and control space is important. When state variables and control variables are continuous, dynamic programming models can only be computed approximately.

Since the numerical routine cannot handle a continuous space, we have to approximate this continuous space by a discrete one. The set of discretized values is called 'grid'. While the approximation is visibly better if the state and control space are very fine (i.e. have many points), this can be costly in terms of computation time. Thus there is a trade-off involved. In practice, the *grid* is usually uniform, with the distance between two consecutive elements being constant.

---

[9] Please see Appendix B to see the demonstration of this equation.

To solve a finite horizon problem, one possible numerical method widely used is the 'backward induction algorithm'. The idea of backward induction is to solve the problem from the end and working backwards towards the initial period. Starting at the last time period, compute the value function for each possible state $x \in X$, and then step back another time period. This way, at time period t we have already computed $V_{t+1}(x_{t+1})$. The crucial limitation is the requirement that we compute the value function $V_t(x_t)$ for all states $x \in X$, this is what we called the *curse of dimensionality*.

Outline of the *backward induction algorithm*:

Step 1: Determine the value function for t=T for all $x_T \in X$.

Set t=T-1.

Step 2: For every $x \in X$ choose $c_t \in X$ which maximizes (or minimizes, depending on the problem objective) the value function in t, i.e., evaluate Equation 4 for all $x_t \in X$.

Step 3. If t>1, decrement t and return to step 2. Else, stop.

Chapter 5

The Model

5.1 The Optimal Execution Problem

Consider an investor seeking to purchase a large block of S shares of some stock, within a fixed finite number of periods T[10]. Given a set of price dynamics that capture price impact (i.e., an individual trade's execution price as a function of the shares traded and other "state" variables), the investor wants to find the optimal sequence of trades (as a function of the state variables) that minimizes the expected execution costs.

As we mentioned before, the short-term demand curves for even the most actively traded equities are not perfectly elastic, hence a market order at time t=1 for the entire block S is clearly not an optimal trading strategy.

Let S_t be the number of shares of the stock acquired in period *t* at prices P_t where *t* = 1, …, *T*, and λ the risk aversion parameter. We can express the investor's objective of minimize the expected execution costs – first term of u(.) – and volatility risk – second term of u(.) – as

$$\min_{\{S_t\}} E_1 \left[\sum_{t=1}^{T} u(P_t S_t)\right] \quad (5)$$

where

$$u(P_t S_t) = P_t S_t + \lambda (P_t S_t)^2 \qquad (6)$$

subject to the constraint

$$\sum_{t=1}^{T} S_t = S \qquad (7)$$

[10] The model in this dissertation is built in a discrete-time way, and the holding period (T) is required to be determined externally. The discrete-time model better fits reality, because a trader could not make sales in a continuous mode. In addition, a time interval might have to be long enough for the restoration of equilibrium. Continuous-time models cannot deal with that. In that respect, it becomes inconsistent with the assumption of the temporary market mechanism described before.

u(.) is also known as 'utility function'. The expected value of the second term of the utility function is the second moment (and not the variance, otherwise we will have to solve an expected value of a variance which is much more complex) of the execution costs which characterizes the volatility.

The parameter λ has a direct financial interpretation. It is already apparent from (6) that λ is a measure of risk-aversion, that is, how much we penalize variance relative to expected cost.

We should also desire to impose a no-sales constraint $S_t \geq 0$, if we don't want to sell stocks as part of a buy-program.

To complete the statement of the problem, we must specify the law of motion for P_t. This includes two distinct components: the dynamics of P_t in the absence of our trade (the trades of others may be causing prices to fluctuate), and the impact that our trade of S_t shares cause on the execution price P_t.

5.2 The Linear-Percentage Temporary Price Impact

To define the state equations we use the 'linear-percentage temporary' (LPT) law of motion of Bertsimas and Lo (1998). Specifically, let the execution price at time t, P_t, be the sum of two components, a no-impact price $\tilde{P}_t$, and the price impact Δ_t caused by purchase[11] a large number of shares:

$$P_t = \tilde{P}_t + \Delta_t \qquad (8)$$

because each trade has a price impact that tends to move the price up for buys and down for sells.

The "no-impact" price is the price that would prevail in the absence of any market impact. A reasonable and observable proxy for the no-impact price is the midpoint of the bid/ask spread, although it can be arbitrary so long as the trade size S_t does not

[11] If we are selling a large block of shares the law of motion will be $P_t = \tilde{P}_t - \Delta_t$ as the selling yields to an increase of the supply and hence to a diminishing of the execution price.

affect it. For convenience, and to ensure non-negative prices, Bertsimas and Lo model the price dynamics of $\tilde{P}_t$ as a geometric Brownian motion:

$$\tilde{P}_t = \tilde{P}_{t-1}\exp(Z_t) \quad (9)$$

Where Z_t is a normal random variable with mean μ_Z and variance σ_Z^2. The exp(.) operator denotes the exponential function. Hence, exp (Z_t) is a Lognormal distribution.

The price impact Δ_t captures the impact of purchasing S_t shares on the transaction prices P_t. Can be defined as a percentage of the no-impact price, $\tilde{P}_t$, and as a linear function of the number of shares and other state variable X_t:

$$\Delta_t = (\theta S_t + \gamma X_t)\tilde{P}_t \quad (10)$$

This specification of the dynamics of P_t has several advantages over other specifications. First, $\tilde{P}_t$ is guaranteed to be nonnegative, and hence P_t (under mild restrictions[12] on Δ_t). Second, separating the transaction price P_t into the no-impact component $\tilde{P}_t$ and the impact component Δ_t turns the trade's price impact temporary. As so, the impact affects the current transaction price but does not affect future prices (observe that the Equation 8 do not depends on last prices P_{t-1}). Third, the percentage price impact increases linearly with the trade size, which is empirically conceivable. The presence of X_t in Equation 10 captures the potential influence of changing market conditions or private information about the security on the price impact Δ_t. Last, Equation 8 implies a natural decomposition of execution costs, decoupling market microstructure effects from price dynamics, which is closely related to the notion of Perold (1988) about *implementation shortfall*. In particular, multiplying Equation 8 by S_t the first term, $\tilde{P}_t S_t$, gives the no-impact cost of execution, the second, $\Delta_t S_t$, is the total impact cost and, therefore, the sum of the two terms gives the actual cost. So the second term is the *implementation shortfall* in executing S_t.

To complete the specification of the state equation, it must be specified the dynamics of the state variable X_t:

$$X_t = \rho X_{t-1} + \eta_t \quad (11)$$

[12] θ, γ and X_t positive.

Where η_t is a white noise (WN) with mean 0 and variance σ_η^2. It is known that information arrival is random. Since information affects market liquidity, the random nature of its arrival will cause liquidity to fluctuate. Because X_t is a autoregressive process with one lag – an AR(1) process – it is possible to capture varying degrees of predictability in information or market conditions.

The parameters θ and γ measure the price impact's sensitivity to trade size and market conditions, respectively.

5.3 The Dynamic-Programming Solution

We use a stochastic dynamic-programming algorithm to solve the optimal-execution problem (see Equation 5). In our context, the state at time t=1, …, T consists of the price $\tilde{P}_{t-1}$ realized at the previous period, the market information at time t, X_t and W_t, the number of shares that remain to be purchased at time t:

$$W_t = W_{t-1} - S_{t-1} \qquad (12)$$

$$W_1 = S \qquad (13)$$

$$W_{T+1} = 0 \qquad (14)$$

The condition $W_{T+1} = 0$ ensures that all S shares are executed by time T.

The state variables summarize all the information which the investor needs in each period t to make his decision regarding the control. The control variable at time t is the number of shares S_t purchased. The randomness is characterized by the random variables $\exp(Z_t)$ and γ_t. The objective is given by Equation 5, $\min_{\{S_t\}} E_1\left[\sum_{t=1}^{T} u(P_t S_t)\right]$, while the law of motion is given by Equation 8, $P_t = \tilde{P}_t + \Delta_t$.

The dynamic programming algorithm is based on the observation that a solution or 'optimal control' $\{S_1^*, S_2^*, \dots, S_T^*\}$ must also be optimal for the remaining program at every intermediate time t. That is, for every t, 1<t<T, the sequence $\{S_t^*, S_{t+1}^*, \dots, S_T^*\}$ must still be optimal for the remaining program $E_t[\sum_{k=t}^{T} u(P_k S_k)]$ where E_t is the conditional expected value on t. This important property is summarized by the

Bellman's Equation (see Chapter 4 for more detail), which relates the optimal value of the objective function in period t to its optimal value in period t+1:

$$V_t(\tilde{P}_{t-1}, X_t, W_t) = \min_{S_t} E_t\left[u(P_t S_t) + V_{t+1}\left(\tilde{P}_t, X_{t+1}, W_{t+1}\right)\right] \quad (15)$$

As with all dynamic-programming solutions, we solve the *Bellman's Equation* recursively, starting at the end. V_T is the optimal value function at the end of our trading horizon, i.e. in period *T*, as a function of the three state variables. By definition,

$$V_T(\tilde{P}_{T-1}, X_T, W_T) = \min_{S_T} E_T[P_T S_T + \lambda P_T^2 S_T^2] \quad (16)$$

$$= \tilde{P}_{T-1} q(1 + \theta W_T + \gamma X_T) W_T + \lambda W_T^2 \tilde{P}_{T-1}^2 (1 + \theta W_T + \gamma X_T)^2 (r + q^2) \quad (17)$$

where q is the mean of a $Lognormal(\mu_Z, \sigma_Z^2)$:

$$q = E_T[\exp(Z_T)] = \exp\left(\mu_Z + \frac{\sigma_Z^2}{2}\right) \quad (18)$$

and r is the variance of a $Lognormal(\mu_Z, \sigma_Z^2)$:

$$r = Var_T[\exp(Z_T)] = [\exp(\sigma_Z^2) - 1] \times \exp(2\mu_Z + \sigma_Z^2) \quad (19)$$

Because this is the last period and W_{T+1} must be set to zero, the remaining order W_T must be executed. Thus, the optimal trade size S_T^* is equal to W_T. Equation 17 shows that the optimal value function is linear-quadratic in X_T and linear-quartic in W_T.

In the next-to-last period T-1, derive a closed-form analytical solution of the *Bellman's Equation* is less trivial, to obtain the optimal trade S_{T-1}^* and hence the optimal-value function V_{t-1} analytically implies to solve a third degree equation, so the optimal control can be derived numerically, using *backward induction algorithm* and applying *Bellman's principle of optimality* (see Chapter 4), as a function of the state variables that characterize the information that the investor must have to make his decision in each period. For t=T-1 up to t=1,

$$V_t(\tilde{P}_{t-1}, X_t, W_t) = \min_{S_t} E_t\left[(P_t S_t + \lambda P_t^2 S_t^2) + V_{t+1}\left(\tilde{P}_t, X_{t+1}, W_{t+1}\right)\right] \quad (20)$$

Where

$$E_t[P_t S_t + \lambda P_t^2 S_t^2] = \tilde{P}_{t-1} q(1 + \theta S_t + \gamma X_T) S_t + \lambda S_t^2 \tilde{P}_{t-1}^2 (1 + \theta S_t + \gamma X_T)^2 (r + q^2) \quad (21)$$

And, assuming independence of $\tilde{P}_t$ and X_{t+1} ($W_{t+1} = W_t - S_t$ is constant in period t),

$$E_t\left[V_{t+1}\left(\tilde{P}_t, X_{t+1}, W_{t+1}\right)\right] = \int_{\exp\,(log(\tilde{P}_{t-1})+\mu_Z-3\sigma_Z)}^{\exp\,(log(\tilde{P}_{t-1})+\mu_Z+3\sigma_Z)} \int_{\rho X_t+\mu_\eta-3\sigma_\eta}^{\rho X_t+\mu_\eta+3\sigma_\eta} V_{t+1}\left(\tilde{P}_t, X_{t+1}, W_t - S_t\right) \times f(\tilde{P}_t) \times g(X_{t+1}) d\tilde{P}_t dX_{t+1} \quad (22)$$

Where log(.) is the natural logarithm, f is the probability density function (PDF) of $\tilde{P}_t$ and g is the PDF of X_{t+1}:

$$f \sim Lognormal(log\tilde{P}_{t-1} + \mu_Z, \sigma_Z^2) \qquad (23)$$

$$g \sim Normal(\rho X_t + \mu_\eta, \sigma_\eta^2) \qquad (24)$$

Please see the proofs in Appendix B.

5.4 The Execution Costs

The Execution Costs (EC) of purchasing S shares of a stock with an initial price P_0 over T periods, in cents/share above the no-impact cost $P_0 S$ is given by:

$$EC = \frac{V_1(\tilde{P}_0, X_1, W_1) - P_0 S}{S} \times 100 \qquad (25)$$

Another measure of the execution costs in percentage is obtained by dividing the transaction cost by the no-impact cost:

$$EC\% = \frac{V_1(\tilde{P}_0, X_1, W_1) - P_0 S}{P_0 S} \times 100\% \qquad (26)$$

5.5 The Parameter Estimation

The parameter estimation procedure consists of three steps.

5.5.1 No-impact Price

First, we estimate the parameters μ_Z and σ_Z^2 of the no-impact price dynamics (see Equation 9). Given the geometric-Brownian motion specification, we know that the returns Z_t are independently and identically distributed (IID) normal random variables:

$$Z_t = log\left(\frac{\tilde{P}_t}{\tilde{P}_{t-1}}\right) \sim N(\mu_Z, \sigma_Z^2) \qquad (27)$$

and $N(\mu_Z, \sigma_Z^2)$ is the normal distribution with mean μ_Z and variance σ_Z^2.

The no-impact price is taken to be the midpoint of the prevailing (best) bid and ask prices at time t:

$$\tilde{P}_t = \frac{\tilde{P}_t^{bid} + \tilde{P}_t^{ask}}{2} \qquad (28)$$

where $\tilde{P}_t^{bid}$ and $\tilde{P}_t^{ask}$ are the best bid and ask prices at time t (ask>bid).

5.5.2 Market Information

The second task is to estimate the parameters of the market-information process in Equation 11. The variable X_t captures the potential impact of changing market conditions or private information about the security. X_t can be described by the returns of a main index, e.g., the S&P 500 Index[13], which behaves accordingly to market conditions.

We rescale the returns by subtracting out the mean and dividing by standard deviation. This yields to a zero-mean, unit standard-deviation process:

$$\tilde{X}_t = \frac{X_t - \mu_X}{\sigma_X} \qquad (29)$$

Once X_t is an AR(1) process and assuming $|\rho| < 1$,[14] we can rewrite the standardized log returns as

$$\tilde{X}_t = \rho\tilde{X}_{t-1} + \eta_t \qquad (30)$$

The maximum likelihood estimator of the AR(1) coefficient, ρ, is

$$\rho = \frac{\frac{1}{T_1}\sum_{t=2}^{T_1} \tilde{X}_t\tilde{X}_{t-1}}{\frac{1}{T_2}\sum_{t=1}^{T_2} \tilde{X}_t^2} \qquad (31)$$

[13] The S&P 500 Index measures changes in stock market conditions based on the average performance of 500 widely held common stocks.

[14] This condition yields to a stationary AR(1) process.

The constants, T_1 and T_2, are the number of observations that are included in calculating the numerator and denominator.

ρ is a measure of serial correlation – or autocorrelation – the correlation of a variable with itself over successive time intervals. If serial correlation exists implies that the correlation between $\tilde{X}_t$ and $\tilde{X}_{t-1}$ is not zero. Serial correlation determines how well the past price of a security predicts the future price.

Given ρ, the maximum-likelihood estimator for the standard deviation of η_t is

$$\sigma_\eta = \sqrt{1-\rho^2} \qquad (32)$$

The parameters ρ and σ_η fully characterize the AR(1) process that describes the market-information.

5.5.3 Price-impact Equation

Our final assignment is to estimate the parameters θ and γ of the price-impact equation, rewriting Equation 8 (and using Equation 10) we obtain:

$$\frac{P_t-\tilde{P}_t}{\tilde{P}_t} = \theta S_t + \gamma X_t \qquad (33)$$

This expression shows that the percentage price impact is a linear function of the volume we intend to trade in the security, and the market information. As before, we form the no impact price, $\tilde{P}_t$, as the average of the bid and ask (see Equation 28) and then construct the dependent variable,

$$\frac{P_t-\tilde{P}_t}{\tilde{P}_t} \qquad (34)$$

for each trade.

The parameters θ and γ can be estimated by a linear regression.

5.5.4 Risk Aversion Parameter

The risk aversion parameter (λ) indicates the weight placed on the price volatility component (which the portfolio is exposed during liquidation) compared to cost.
The risk aversion parameter can be either positive or negative.

If lambda is greater than zero, it would be chosen by a risk-averse trader who wishes to sell quickly to reduce exposure to volatility risk, despite the trading costs incurred in doing so.
When lambda is equal to zero, we'll call this the naive strategy, since it represents the optimal strategy corresponding to simply minimizing expected transaction costs without regard to variance (like the Bertsimas and Lo approach). Almgren and Chriss demonstrate that in a certain sense, this is never an optimal strategy, because one can obtain substantial reductions in variance for a relatively small increase in transaction costs.
When lambda is negative it would be chosen only by a trader who likes risk. He postpones selling, thus incurring lower expected trading costs but higher variance during the extended period that he holds the security.

Chapter 6

Empirical Study

We now implement the optimal execution strategy for a particular security. Specifically, we estimate the parameters of the linear-percentage model described in the previous chapter. We then compare the best-execution strategy with the “naive” strategy which corresponds to simply minimizing expected transaction costs without regard to volatility risk.

6.1 Data

In order to estimate the parameters of the no-impact price dynamics and the price-impact equation it was collected market data from NASDAQ OMX[15]. More precisely, the data contains daily last price (see Figure 4), best bid and ask prices and traded number of shares from 1 January 2009 to 31 December 2009 (1 year) from a stock of a Baltic Bank (AB Ukio bankas) [16]. The currency is Lithuanian Litas (LTL).

ISIN	LT0000102352
Name	UKB1L
Market	BALTIC MAIN LIST
Nominal value	1.00 LTL
Total number of securities	245,824,000
Core business	Banking activities

Table 2 – Security information

15 http://www.nasdaqomxbaltic.com/market/?instrument=LT0000102352&list=2&currency=EUR&pg=details&tab=historical&lang=en

16 The choice of the security it concerns to availability of bid and ask prices.

Figure 4 – Last price (in LTL) of Ukio Bankas Security from January to December 2009

To estimate the parameters of the market information process it was collected daily close price information of the S&P 500 INDEX (currency in U.S. dollar) for the same time period (1 January 2009 to 31 December 2009) from Yahoo Finance[17] as shown in the following figure:

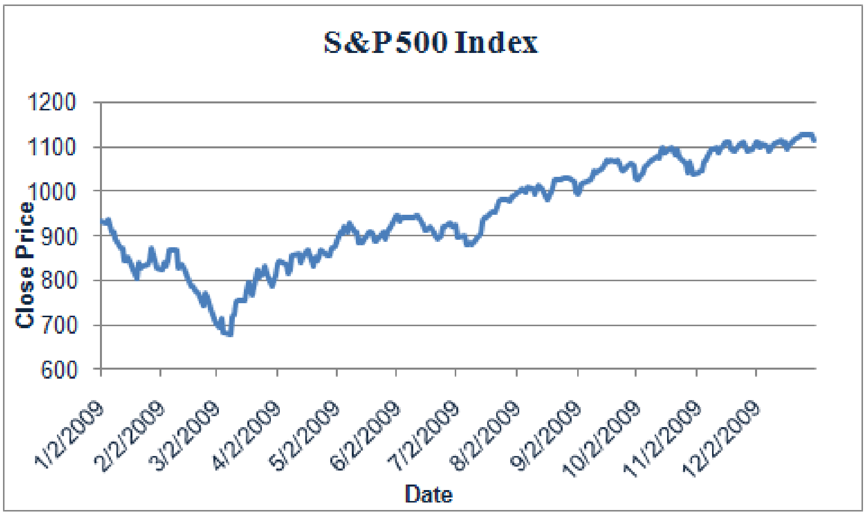

Figure 5 – Close price (in USD) of S&P500 Index from January to December 2009

[17] http://finance.yahoo.com/q/hp?s=%5EGSPC

6.2 Parameter Estimation

6.2.1 No-impact Price

After collecting data from NASDAQ OMX at every day over the course of the 2009 trading year we calculate the midpoint of the bid and ask prices (as shown in Figure 6) to construct the no impact price, $\tilde{P}_t$.

Figure 6 – Mid price of Ukio Bankas Security from January to December 2009

We then form log returns (please see Figure 7) according to Equation 27 and this gives us a sample of 247 observations of Z_t during the 2009 trading year from which we can estimate μ_Z (mean) and σ_Z (standard deviation) in the standard way.

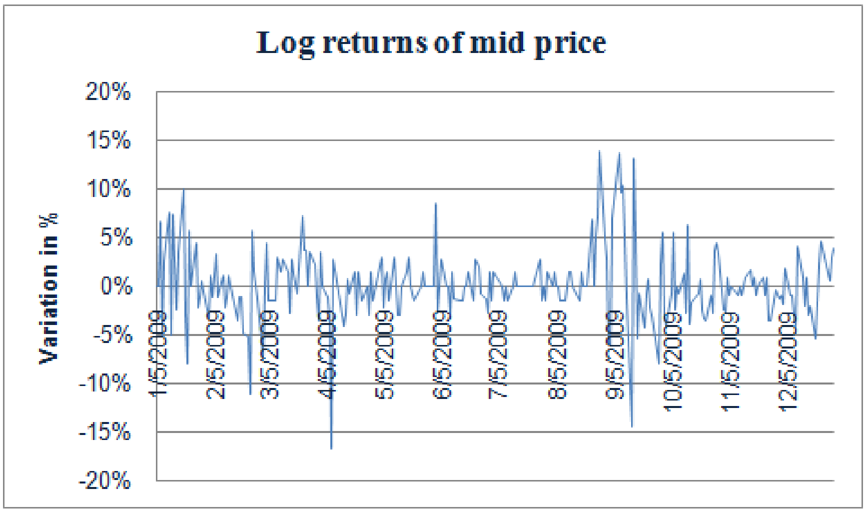

Figure 7 – Mid price log returns of Ukio Bankas Security from January to December 2009

Table 3 summarizes the results expressed in percent. The drift (μ_Z) and volatility (σ_Z) estimates are consistent with intuition and fit well with other studies.

Parameter	Estimation	Annualized
σ_Z	3.7%	58.8%
μ_Z	0.1%	36.6%

Table 3 – Estimated drift and volatility of no impact price

Note: The volatility σ_Z was estimated also using an exponentially weighted moving average (EWMA), such that a higher weight is placed on more recent observations. The daily result was 2.8%.

6.2.2 Market Information

In our empirical exercise, the market information X_t denotes the returns on the S&P 500 index, a common factor that influences the prices of most securities.

Using the S&P 500 data from 1 January to 31 December 2009, we construct the daily returns X_t, yielding to 251 observations, as shown in the next figure:

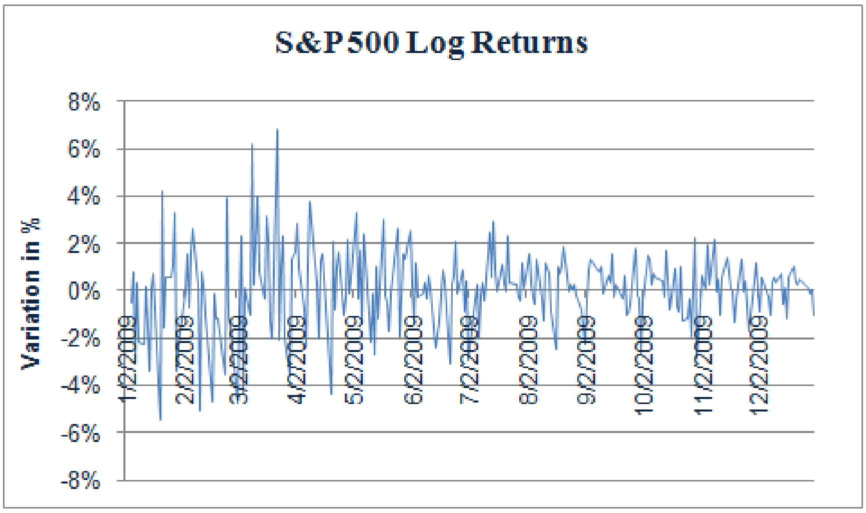

Figure 8 – Log returns of S&P500 Index from January to December 2009

Rescaling the returns subtracting the mean (0.1%) and dividing by the standard deviation (1.7%) we can obtain the maximum likelihood estimator of the AR(1) coefficient ρ which is -0.1158. (Not unexpected the level of serial correlation in the S&P 500 index to be quite low. If not, profitable trading strategies would be possible due to the predictability of index returns). Hence, following the Equation 32, σ_η is 0.9933.

6.2.3 Price Impact Equation

The NASDAQ database provides one independent variable – namely, the volume of the stock, S_t. The other independent variable is the returns of the S&P 500 index, X_t. The dependent variable represents the percentage price impact (see Equation 34).

We now have a complete set of data with which to estimate the parameters of the price impact model. We performed the regression in Excel; it contained no intercept term because the price impact should be zero if no stocks are being traded. Table 4 summarizes the regression:

Sample size		242
R Square		0.058
Return S&P (X_t)	Coefficient (γ)	8×10^{-4}
	t-statistic	(0.03)
No of Shares (S_t)	Coefficient (θ)	3×10^{-9}
	t-statistic	(3.85)

Table 4 – Regression output

R^2 indicates that the regression have a low explanatory power although bear in mind that, because of unavailability of information, we have omitted many other variables that proprietary traders and other professional portfolio managers have at their disposal. The stock price-impact term, that is, θ, has *t*-statistic significant at the 5% level. This term should be the most dominant in determining price impact, and our regression confirms this conjecture.

The fundamental assumption in linear regression is that the error term ε_t has mean zero and is independent and identically distributed [E(ε_t) = 0, Var(ε_t) = σ^2, and E($\varepsilon_t \varepsilon_{t-1}$) = 0]. We have performed diagnostics on the residuals to test for the presence of heteroscedasticity (not identical distribution of error terms, i.e., variance of errors not constant) and autocorrelation (dependence of error terms). The Durbin-Watson test indicated low levels of positive serial correlation, with statistics equal to 1.79.

The White test indicated a weak presence of heteroscedasticity, because the *p*-value is low.

For more details about these tests please see Appendix C.

6.3 Numerical Dynamic Algorithm

In order to program the *backward induction algorithm*, after choosing a functional form for the utility function and estimating the parameters the next step is to discretize the state and control variable.

We have to define the space spanned by the state and the control variables. The computer cannot literally handle a continuous state space, so we have to approximate this continuous space by a discrete one. As we mentioned before, while the approximation is clearly better if the state space is very fine (i.e. has many points), this can be costly in terms of computation time. Thus there is a trade-off involved. The natural state space W_t is given by a uniform *grid* {0, …, S}, where S is the total number of shares, with the distance between two consecutive elements being constant. Let n_W be the number of elements in the state space[18]. The control variable, S_t, also takes values between 0 and S, with length n_W.

We now review in more detail how the iteration is done in practice. Starting on t=T up to t=1, at each iteration, the values of the *Bellman's Equation* $v_j(W)$ are stored in a n_Wx1 matrix:

[18] A possible choice for n_W is S+1, the case when the distance between two consecutive elements is equal to one.

$$V = \begin{bmatrix} v_j(W_1) \\ \vdots \\ v_j(W_s) \\ \vdots \\ v_j(W_{n_W}) \end{bmatrix}$$

To compute v_{j+1}, we start by choosing a particular size for the total amount of shares at the start of the period, W_T. We then search among all the points in the control space for the one that minimizes the expected utility (see Equation 20). Let's denote it S_i^*.

This involves finding next period's value. Once we have calculated the new value for $v_{j+1}(W_s)$, we can proceed to compute similarly the value $v_{j+1}(.)$ for other sizes of the portfolio at the start of the period. These new values are then stacked in V. Please see Appendix D to see the form of the value function.

The *backward induction algorithm* was written in MATLAB Code, please see Appendix E for the full algorithm.

6.4 Results and Discussion

Having calibrated the price-impact equation we now analyse the results of the optimal-execution strategy through the *backward induction algorithm*. As we saw in section 6.2 the estimated parameters are:

Parameter	Estimation
θ	3%
γ	0.0008
ρ	-0.12
σ_Z	0.037

Table 5 – Estimated parameters of the linear-percentage temporary price impact model

Specifically, the goal is to minimize the expected execution costs of a 10,000,000-share purchase for a stock currently trading at $P_0 = 1.03\ LTL$. The parameter θ is calibrated to yield a percentage price impact of 3% for a 10,000,000-share trade. [19]

To gauge the sensitivity of execution costs to the model's parameters, we vary the price impact parameters, θ (sensitivity of price impact to trade size) and γ (sensitivity of price impact to market conditions), the autocorrelation coefficient, ρ, the no-impact price volatility, σ_Z^2, the time horizon, T, and the risk aversion parameter, λ. We modify the parameters by scaling by a constant.

For θ the scaling factors are 1, 2 and 4. Table 6, reports the estimated execution cost in cents per share, for our optimal-execution strategy with λ=0.1 ("Optimal") and for the model which ignores the volatility term, i.e, with λ=0 ("Naïve").

θ	Optimal	Naïve	Difference
3%	14.09	2.62	-11.46
6%	14.47	2.94	-11.53
12%	15.24	3.57	-11.67

Table 6 – Estimated execution costs (in cents per share) for the optimal execution strategy and for the naïve strategy for different values of the price impact parameter θ

We can observe that when the sensitivity of price impact to trade size raises the transaction costs also augment, once the transaction costs are an increasing function of the number of shares purchased.

[19] To develop some intuition for the coefficients, consider the estimated price impact for $\hat{\theta}$ and $\hat{\gamma}$ caused by trading in the stock UKB1L, which is 3% (in percentage of the number of shares) and 8×10^{-4}, respectively. Assuming T=1 (implies variance=0), if we traded a 10,000,000-share block of UKB1L stock at its beginning-of-year price of 1.03 LTL with no impact, our total cost would be $P_0 S$, i.e. 10,000,000 × 1.03 LTL = 10,300,000 LTL. But according to Table 5, the full-impact cost would be

$$P_t S = 10{,}000{,}000 \times (\tilde{P}_t + \Delta_t) = 10{,}000{,}000 \times (1.03 + \Delta_t) = 10{,}609{,}000\ LTL$$

Where $\Delta_t = \theta S_t \times \tilde{P}_t = 3\% \times \tilde{P}_t = 0.0309$, which implies a price impact of approximately 3 cents per share (ignoring the other factors in the regression, i.e., the γX_t term).

The parameter γ is scaled by 0, 0.5, 1, 2 and 4, respectively. The results are the following:

γ	Optimal	Naïve	Difference
0	14.09	2.63	-11.46
0.0004	14.09	2.63	-11.46
0.0008	14.09	2.62	-11.46
0.0016	14.08	2.62	-11.46
0.0032	14.06	2.60	-11.46

Table 7 – Estimated execution costs (in cents per share) for the optimal execution strategy and for the naïve strategy for different values of the price impact parameter γ

When γ increases, implying that information X_t has a larger effect on price impact Δ_t, the best-execution strategy (and the naïve strategy) performs even better. Our strategy optimally uses information so as to trade when trading is least expensive. For sufficiently significant pieces of information, trading can be quite profitable.

ρ	Optimal	Naïve	Difference
-0.12	14.09	2.62	-11.46
0	14.11	2.64	-11.46
0.12	14.13	2.66	-11.47

Table 8 – Estimated execution costs (in cents per share) for the optimal execution strategy and for the naïve strategy for different values of the price impact parameter ρ

The term X_t indicates the presence of serially correlated information. If $\rho < 0$ so that X_t exhibits reversals[20], a positive realization of X_t decreases the number of shares purchased, *ceteris paribus*: it is more expensive to trade in period t and X_t is likely to reverse next period making it less expensive to trade then, hence it is optimal to trade less now.

When $\rho = 0$, implies that X_t is unforecastable, and while X_t still has an impact on the current execution price, observing X_t tells us nothing about expected future execution prices hence it can no longer affect the best-execution strategy.

[20] A change in the direction of a price trend.

Alternatively, if ρ > 0, the execution costs rise, this may seem somewhat counterintuitive at first because positive realizations of X_t necessarily increases the execution price now, P_t, by γX_t, so why trade more? The answer may be found in the fact that information X_t is positively serially correlated, hence $X_t > 0$ implies that future realizations are likely to be positive which, in turn, implies higher trading costs for several periods thereafter[21] (on average). Therefore, although a positive X_t makes it more costly to purchase shares in period t, this additional cost is more than offset by the sequence of expected future price increases that arise from positively serially-correlated information[22].

Then, we modify the price volatility of the mid price returns by scaling the variances by a constant of 0.25, 0.5 and 1, respectively:

σ_Z	Optimal	Naïve	Difference
0.009	13.11	1.84	-11.26
0.019	14.08	2.64	-11.44
0.037	14.09	2.62	-11.46

Table 9 – Estimated execution costs (in cents per share) for the optimal execution strategy and for the naïve strategy for different values of the daily standard deviation σ_Z

Increasing volatility seems to increase execution costs slightly.

The following table shows the trading costs results for different time periods (in days):

T	Optimal	Naïve	Difference
5	33.14	5.00	-28.14
10	15.45	3.75	-11.70
15	14.42	2.89	-11.52
20	14.09	2.62	-11.46

Table 10 – Estimated execution costs (in cents per share) for the optimal execution strategy and for the naïve strategy for different values of the number of periods T (in days)

[21] In particular, a one-unit increase in X_t implies an expected increase in P_{t+1} of $\gamma\rho$, an expected increase in P_{t+2} of $\gamma\rho^2$, and so on.

[22] Trading 'into' expected price increases can be more economical than the natural tendency to wait for a more favorable price.

We can observe that as T increases, execution costs fall. Because we can spread the trading over more time periods, and because we have the flexibility to be more patient and wait for particularly opportune times to trade, expected costs decline. Also, as S_t decreases over time, execution costs also decrease.

Finally, we estimate the execution costs for different values of the risk aversion parameter.

λ	Costs	λ	Costs
0 (Naïve)	2.62	0 (Naïve)	2.62
0.01	3.77	-0.01	1.47
0.1	14.09	-0.1	-8.84
0.5	59.93	-0.5	-54.68
1	117.24	-1	-111.98

Table 11 – Estimated execution costs (in cents per share) for the optimal execution strategy and for the naïve strategy for different values of the risk aversion parameter λ

As expected, the execution costs increase as λ increases. This is because when λ is greater than zero, it would be chosen by a risk-averse trader who wishes to sell quickly to reduce exposure to volatility risk, despite the trading costs incurred in doing so. When λ is equal to zero (naive strategy – Bertsimas and Lo approach) corresponds to simply minimizing expected transaction costs without regard to variance. When λ is negative corresponds to a trader who postpones selling, thus incurring lower expected trading costs but higher variance during the extended period that he holds the security.

In all, the naive strategy underestimates the execution costs because this strategy disregards the volatility risk.

Chapter 7

Conclusion

Market liquidity risk is, in some aspect, more complex than other risk types, because it is influenced by the actions of the trader, who has the choice whether to buy or sell a position more or less quickly. Hence, the risk exposure and cost of trading vary according to the chosen strategy. With the advent of more accurate models and measures of execution costs, and given the remarkable increase in institutional trading in past years, the optimal control of execution costs has assumed a huge importance.

This dissertation studied the optimal behaviour of a trader who wishes to buy (or sell) a given quantity of a security within a certain time period. His trades affect current and future prices of the security. He therefore breaks up his trades into a sequence of smaller orders. We have argued that this is a dynamic optimization problem because trading takes time, the demand for financial securities is not perfectly elastic, and the price impact of current trades, even small trades, can affect the course of future prices.

Unlike some studies, like Bertsimas and Lo (1998), which miss an important component of liquidity, the proposed approach accounts for both, average cost and, importantly, risk (given by the second moment of the costs). One of the main features of the model presented in this dissertation with one source of cost (price impact) and one source of risk (price volatility) is that there is a clear cost-risk trade-off depending on the trader’s level of risk aversion.

The central goal of our analysis has been the construction of a *backward induction algorithm* using numerical stochastic dynamic programming, in order to derive an optimal trading strategy to execute a large order of a given security that minimize the expected value and the second moment of cost execution - called best execution strategies. The best execution strategy varies through time as a function of state variables which measure market conditions and the remaining shares to be executed, accordingly with the price-impact model of Bertsimas and Lo. This thesis also demonstrates that expected cost and risk components of transaction costs can be estimated from detailed stock data.

Several conclusions of practical importance followed from this analysis:

- The transaction costs are an increasing function of the number of shares traded;
- When the weight of market information rises transaction costs diminish, once access to information and market conditions imply a speculation about market tendency and hence the investor can profit from that;
- Higher risk aversion or price volatility cause a higher trading costs;
- Higher price volatility cause upper trading costs;
- As time increases, execution costs fall, because when we spread the trading over more time periods, we have the flexibility to be more patient and wait for particularly opportune times to trade. Also, as the quantity transacted decreases over time, execution costs also decrease.

Optimal liquidation strategies are appealing to risk management, because the time component of liquidity is explicitly modelled. Knowing the optimal strategy, several interesting measures can be derived for market and funding liquidity risk management purposes.

Despite the progress during recent years, solutions for liquidity risk measurement still are not as clear as for market risk measurement. Liquidity is a multidimensional problem that is hard to comprise in one measure. Risks associated with liquidity depend to a large degree on individual factors like time to liquidation, price impact and risk aversion. Nevertheless, it is worthwhile to put considerable effort in this area of research, as a more accurate estimation of risk leads to more efficient use of capital and an increased awareness of potential troubles.

Future Research

In order to deep and give continuity to the exposed work, future research can go in three main directions:

Time-varying coefficients: Optimal trading strategies can be further improved to take into account time varying liquidity by applying a time varying price impact coefficient and also the volatility and the expected drift can be time-dependent as long as their values are known at the start of liquidation.

A resulting optimal strategy would better reflect actual liquidity situations found in the market. As a general rule, there should be more trading in periods of high liquidity and less trading in low liquidity periods.

Finding the optimal strategy entails solving a linear system of size equal to the number of time periods (times the number of stocks, for a portfolio problem). One example in which this is useful is if the price is expected to jump either up or down on a known future date (e.g., an earnings announcement), as long as we have a good estimate of the expected size of this jump.

Portfolio: common to most approaches for liquidity risk assessment is the focus on a single asset. If any of the models is extended to the multiple asset portfolio case, the calculation is usually complicated and poses a high demand on computational power. For a multi-asset portfolio that needs to be liquidated not only the variance-covariance but also cross-impacts should be taken into account.

Nonlinear cost functions: solving a non-quadratic optimization problem; the difficulty of this problem depends on the specific functional form chosen.

We hope that these extensions will lead to further useful insights.

References

Adda, J. and Cooper, R.; *Dynamic economics: quantitative methods and applications*; MIT Press; 2003.

Alfonsi, A. and Schied, A.; *Optimal execution and absence of price manipulations in limit order book models*; 2009.

Almgren, R. and Chriss, N.; *Optimal Execution of Portfolio Transactions*; The Journal of Risk; 2000; 3; pp. 5–39.

Almgren, R. and Chriss, N.; *Value under liquidation*; Risk; 1999; 12; pp. 61-63.

Bangia, A., et al.; *Modeling Liquidity Risk; With Implications for Traditional Market Risk Measurement and Management*; Risk; 1999; 12; pp. 68–73.

Basel Committee on Banking Supervision; *Liquidity Risk: Management and Supervisory Challenges*; February 2008.

Basel Committee on Banking Supervision; *Principles for Sound Liquidity Risk Management and Supervision*; September 2008.

Bellman, R. E. and Dreyfus, S. E.; *Applied dynamic programming*; Princeton University Press, 1962.

Bertsimas, D. and Lo, A. W.; *Optimal Control of Execution Costs*; Journal of Financial Markets; 1998; 1; pp. 1–50.

Bertsimas, D., Lo, A. W. and Hummel, P.; *Optimal control of execution costs for portfolios*; Computing in Science & Engineering; 1999; 1; pp. 40–53.

Bervas, A.; *Market Liquidity and its Incorporation into Risk Management*; Banque de France; Financial Stability Review; 2006; No 8; pp. 63-79.

Bessis, J.; *Risk Management in Banking*; John Wiley & Sons Ltd; 2002.

Brunnermeier, M. K. and Pedersen, L. H.; *Market Liquidity and Funding Liquidity*; NBER Working Paper 12939; 2007; National Bureau of Economic Research.

Committee of European Banking Supervisors; *Second Part of CEBS's Technical Advice to the European Commission on Liquidity Risk Management*; 2008.

Deutsche Bundesbank; *Liquidity risk management at credit institutions*; Monthly Report; September 2008; pp. 57 – 71.

Duttweiler, Rudolf, W.; *Managing Liquidity in Banks: A Top Down Approach*; Wiley Finance; 2009.

Engle, R., Ferstenberg, R. And Russell, J.; *Measuring and Modeling Execution Cost and Risk*; January 2008.

Fragnière, E., Tuchschmid, S. N. and Zhang, Q.; *Measuring Liquidity Risk: The Estimation of Liquidity Adjusted Value at Risk*; 2009.

Glantz, M. and Kissell, R.; *Optimal Trading Strategies: Quantitative Approaches for Managing Market Impact and Trading Risk*; 2003; New York; Amacom.

Hisata, Y. and Yamai, Y.; *Research Toward the Practical Application of Liquidity Risk Evaluation Methods*; Monetary and Economic Studies; 2000; Vol.18; p.83-127.

Huberman, G. and Stanzl,W.; *Optimal Liquidity Trading*; Review of Finance; 2005; 9; pp. 165–200.

Institute of International Finance; *Principles of Liquidity Risk* Management; 2007.

Jorion, P.; *Value at Risk: The New Benchmark for Managing Financial Risk*; New York; 2001; McGraw-Hill.

Judd, Kenneth L.; *Numerical methods in economics*; MIT Press, 1998

Ljungqvist, Lars and Sargent, Thomas J.; *Recursive macroeconomic theory*; MIT Press, c2004.

Malz, M. Allan; *Liquidity risk: current research and practice*; RiskMetrics Group.

Matz, Leonard and Neu, Peter; *Liquidity Risk Measurement and Management: A Practitioner's Guide to Global Best Practices*; Wiley Finance; 2007.

Obizhaeva, A. and Wang, J.; *Optimal Trading Strategy and Supply/Demand Dynamics*; NBER Working Paper 11444; National Bureau of Economic Research; 2005.

Ross, Sheldon M.; *Introduction to stochastic dynamic programming*; San Diego; Academic Press; c1995

Sargent, Thomas J.; *Dynamic macroeconomic theory*; Harvard University Press; 1987.

Schied, A. and Schoneborn, T.; *Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets;* 2008.

Appendix A

Other Liquidity Risk Types

Call liquidity risk (also called withdrawal liquidity risk) relates to both assets and liabilities. Is the risk that more credit lines will be drawn or more deposits withdrawn than expected.

Term liquidity risk refers to an unexpected prolongation of the capital commitment period in lending transactions or payments deviation from contractual conditions (e.g. unexpected delays in repayments).

Contingent liquidity risk is the likelihood that an institution will be called upon to provide liquidity unexpectedly, possibly at a time when it is already under stress.

Appendix B

Proofs

Equation 4: At time t=1, if we decompose the sum and apply the principle of DP, we get

$$V_1(x_1) = E_1\left[\sum_{t=1}^{T} u(c_t^*, x_t)\right] =$$

$$= \max_{\{c_t\}_{t=1}^{T-1}} E_1\left[\sum_{t=1}^{T} u(c_t, x_t)\right] =$$

$$= \max_{\{c_t\}_{t=1}^{T-1}} E_1\left[u(c_1, x_1) + \sum_{t=2}^{T} u(c_t^*, x_t)\right] =$$

$$= \max_{c_1}\left\{u(c_1, x_1) + \max_{\{c_t\}_{t=2}^{T-1}} E_1\left[\sum_{t=2}^{T-1} u(c_t^*, x_t)\right]\right\} =$$

$$= \max_{c_1}\left\{u(c_1, x_1) + \max_{\{c_t\}_{t=2}^{T-1}} E_1\left[E_2\left[\sum_{t=2}^{T-1} u(c_t^*, x_t)\right]\right]\right\}$$

In the last step we use the law of iterated expectations: $E_t[E_s(X)] = E_t(X)$ if $s \geq t$.

Therefore we obtain,

$$V_1(x_1) = \max_{c_1}\{u(c_1, x_1) + E_1[V_2(x_2)]\}$$

The same idea can be extended to any $1 \leq t \leq T$. Then, for time t we get the stochastic *Bellman's Equation* or value function:

$$V_t(x_t) = \max_{c_t}\{u(c_t, x_t) + E_t[V_{t+1}(x_{t+1})]\}$$

Equation 22 comes from the application of the expected value to a function. The expected value of an arbitrary function of X, h(X), with respect to the PDF k(x) is given by the inner product of h and k: $E[h(X)] = \int_{-\infty}^{+\infty} h(x)k_X(x)dx$.

Equation 23 defines the distribution of $\tilde{P}_t$ which is given by:

$$\tilde{P}_t = \tilde{P}_{t-1}\exp(Z_t) = \exp\big(\log(\tilde{P}_{t-1})\big) \times \exp(Z_t) = \exp\,(\log(\tilde{P}_{t-1}) + Z_t) =$$
$$\exp\,(Normal(log(\tilde{P}_{t-1}) + \mu_Z, \sigma_Z^2)) = \text{Lognormal}\big(\log\,(\tilde{P}_{t-1}) + \mu_Z, \sigma_Z^2\big)$$

Where $Z_t \sim Normal(\mu_Z, \sigma_Z^2)$ and in time t $\tilde{P}_{t-1}$is a constant, hence also $log(\tilde{P}_{t-1})$.

The domain of f can be derived in the following way:

$$(log(\tilde{P}_{t-1}) + \mu_Z) - 3\sigma_Z < Normal(\log\,(\tilde{P}_{t-1}) + \mu_Z, \sigma_Z^2) < \big(\log\,(\tilde{P}_{t-1}) + \mu_Z\big) + 3\sigma_Z$$

Hence,

$$\exp\,(log(\tilde{P}_{t-1}) + \mu_Z - 3\sigma_Z) < \exp\,(Normal(log(\tilde{P}_{t-1}) + \mu_Z, \sigma_Z^2))$$
$$< \exp\,(log(\tilde{P}_{t-1}) + \mu_Z + 3\sigma_Z)$$

Equation 24 is the PDF of X_{t+1} and can be defined as:

$$X_{t+1} = \rho X_t + \eta_{t+1} \sim Normal(\rho X_t + \mu_\eta, \sigma_\eta^2)$$

Because $\eta_t \sim WN(\mu_\eta, \sigma_\eta^2)$ and ρX_t is constant in time t.

The domain of g corresponds to the domain of a Normal distribution with expected value given by $\rho X_t + \mu_\eta$ and standard deviation σ_η which is given by:

$$\rho X_t + \mu_\eta - 3\sigma_\eta < Normal\big(\rho X_t + \mu_\eta, \sigma_\eta^2\big) < \rho X_t + \mu_\eta + 3\sigma_\eta$$

Note that three times the standard deviation covers 99.97% of the area under the Normal curve.

Appendix C

Durbin Watson and White Tests

One of the widely used tests for detecting the presence of serial correlation is the Durbin Watson (DW) Test (Durbin and Watson 1950). A test that the residuals from a linear regression or multiple linear regression are independent.

When error terms from different (usually adjacent) time periods are correlated, we say that the error term is serially correlated. Serial correlation occurs in time-series studies when the errors associated with a given time period carry over into future time periods.

There are different types of serial correlation. With first-order serial correlation, errors in one time period are correlated directly with errors in the subsequent time period.

The Durbin Watson test is based on the principle that if the errors are autocorrelated, this fact will be revealed through the autocorrelations of the least squares residuals. The null hypothesis is that errors are not autocorrelated, i.e., $H_0: \rho = 0$.

The test statistic is

$$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}$$

Where $e_i = y_i - \hat{y}_i$ and y_i and $\hat{y}_i$ are, respectively, the observed and predicted values of the response variable for individual i. d becomes smaller as the serial correlation increases. Upper and lower critical values, d_U andd_L, have been tabulated for different values k (number of explanatory variables) and n (number of time periods).

$$If\ d < d_L \quad reject\ H_0: \rho = 0$$

$$If\ d > d_U \quad do\ not\ reject\ H_0: \rho = 0$$

$$If\ d_L < d < d_U \quad test\ is\ inconclusive.$$

When successive values of e_i are close to each other, the DW statistic will be low, indicating the presence of positive serial correlation.

d ranges from zero (perfect positive autocorrelation) to 4 (perfect negative autocorrelation). Note that there are two inconclusive areas where the null hypothesis cannot be tested properly.

One of the classical assumptions of the ordinary regression model is that the variance of the errors is constant, or homogeneous, across observations. If this assumption is violated, the errors are said to be "heteroscedastic". When heteroscedasticity is present the parameter estimates are still consistent but they are no longer efficient. Thus, inferences from the standard errors are likely to be misleading.

There are several methods of testing for the presence of heteroscedasticity. The most commonly used is the White's General Test (White 1980).

The White test is computed by finding nR^2 from a regression of e_i^2 on all of the distinct variables in X⊗X, where X is the vector of dependent variables including a constant. This statistic is asymptotically distributed as a Chi-square with k-1 degrees of freedom, where k is the number of regressors, excluding the constant term.

Notice, however, that the White test rejects the null hypothesis of no heteroscedasticity. This implies that the standard errors of the parameter estimates are incorrect and, thus, any inferences derived from them may be misleading. With heteroscedasticity, the OLS (Ordinary Least Squares) estimator is unbiased and consistent but it is not BLUE (the Best Linear Unbiased Estimator à la Gauss-Markov theorem) or asymptotically efficient.

Appendix D

Value Function of the Optimal Execution Strategy and the Naïve Strategy

The value function for the naive strategy in each iteration is linear:

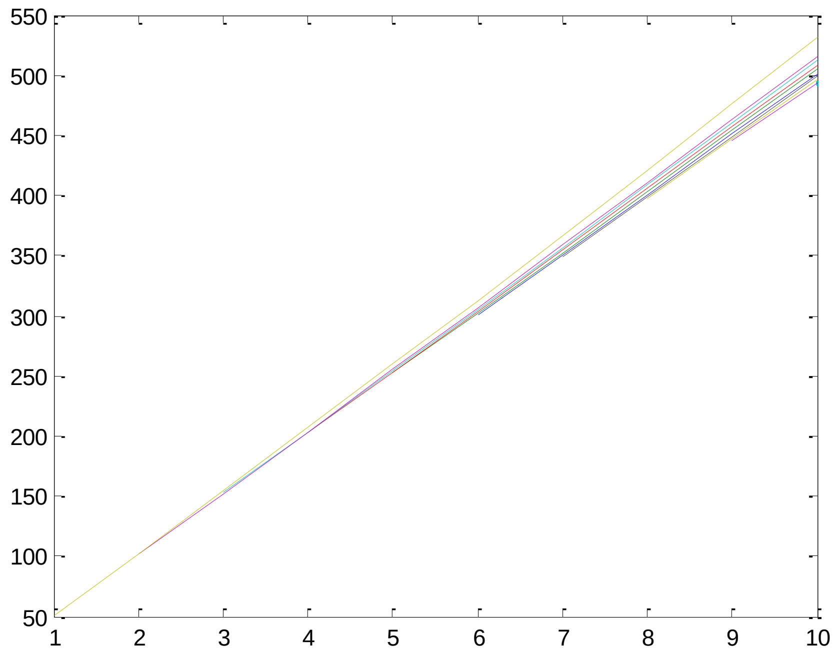

Figure 9 – Value function of the naïve strategy for T=20 and S=10 (considering the estimated parameters)

The value function for the optimal strategy in each iteration is quadratic:

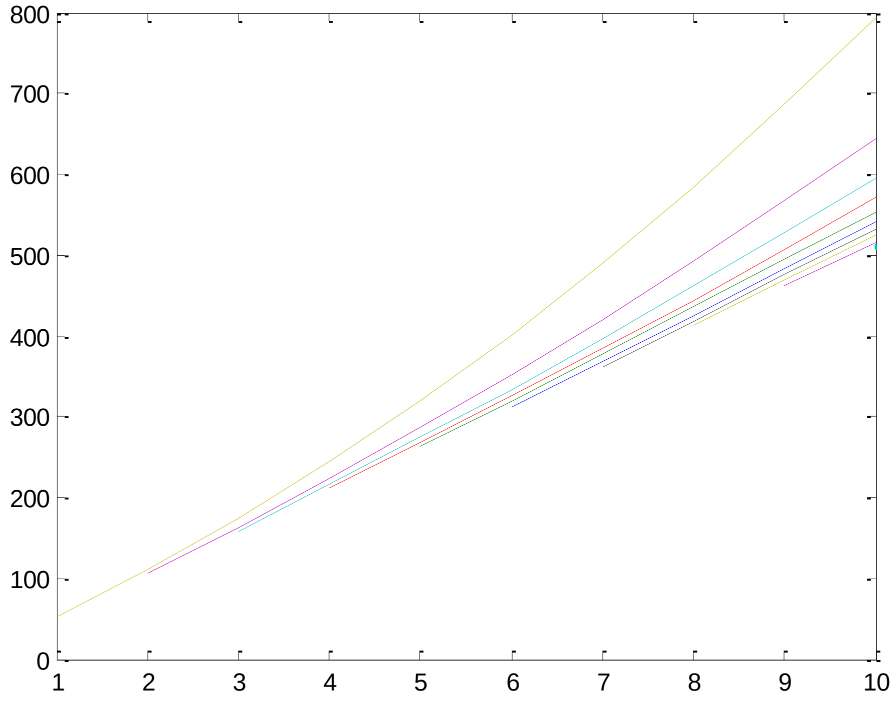

Figure 10 – Value function of the optimal strategy for T=20 and S=10 (considering the estimated parameters)

Appendix E

Numerical Dynamic Programming Algorithm (MATLAB Code)

```
>> TS=10; % total number of shares
T=20; % number of periods
muz=0;
sigmaz=0.037;
theta=0.03/TS;
rho=-0.12;
mueta=0;
sigmaeta=sqrt(1-rho^2);
gamma=0.0008;
lambda=0;

% Define Xt for t=1, ..., T
X(1,1)=0;
for i=2:T
X(1,i)=rho*X(1,i-1)+normrnd(mueta,sigmaeta);
end

q=exp(muz+sigmaz^2/2); % expected value of lognormal
r=exp(2*muz+sigmaz^2)*(exp(sigmaz^2)-1); % variance of lognormal

Z=normrnd(muz,sigmaz,1,T); % Define Zt for t=1, ..., T

% Define Pt for t=1, ..., T
P(1,1)=1.03; % initial price
for i=2:T
P(1,i)=P(1,i-1)*exp(Z(1,i-1));
end

W=1:1:TS; % Define the state space W
S=1:1:TS; % Define the control space S

v=zeros(size(W,2),T)+NaN; % Value function matrix g*T
p=zeros(size(W,2),T); % Policy function matrix g*T


% Initialization:VT
for m=1:size(W,2)
v(m,T)=q*P(1,T)*(1+theta*W(1,m)+gamma*X(1,T))*W(1,m)+lambda*W(1,m)^2* P(1,T)^2*(1+theta*W(1,m)+gamma*X(1,T))^2*(r+q^2);
p(m,T)=W(1,m);
end

% Backward algorithm
for t= T-1:-1:1
for i=2:size(W,2)
j=1;
while j<i
if isnan(v(i-j,t+1))
vaux(j,1)=NaN;
else
F = @(x,y) v(i-j,t+1)*lognpdf(x, log(P(1,t))+muz, sigmaz)* normpdf(y,rho* X(1,t)+mueta,sigmaeta);
ev= dblquad(F,exp(log(P(1,t))+muz-3*sigmaz),exp(log(P(1,t))+muz+3*sigmaz), rho*X(1,t)+mueta-3*sigmaeta, rho*X(1,t)+mueta+3*sigmaeta);
vaux(j,1)=q*P(1,t)*(1+theta*S(1,j)+gamma*X(1,t))*S(1,j)+lambda*P(1,t)^2*(1+theta*S(1,j)+gamma*X(1,t))^2*S(1,j)^2*(r+q^2)+ev;
end
j=j+1;
end
[m,s]=min(vaux);
v(i,t)=m;
p(i,t)=S(1,s);
clear vaux;
end
end

cost=((v(TS,T-TS+1)-P(1,1)*TS)/TS)*100
```